



SAPIENZA
UNIVERSITÀ DI ROMA

FACOLTÀ DI INGEGNERIA

**Dottorato in Ingegneria
dell'Informazione e della Comunicazione
XXI ciclo**

Tesi di Dottorato

**Modelli a poli comuni
per sistemi acustici multicanale**

Coordinatore

Prof. Aurelio Uncini

Candidato

Dott. Gabriele Bunkheila

Anno Accademico 2007/2008

Modelli a poli comuni
per sistemi acustici multicanale

Gabriele Bunkheila

Dicembre 2008

Indice

1	Modelli a Poli Acustici Comuni	3
1.1	Introduzione	3
1.2	La stima dei poli comuni	5
1.2.1	Il metodo HMK	5
1.2.2	Altri metodi per la stima di poli comuni	6
1.3	Giustificazione teorica dei CAP	7
1.3.1	CAP e reti a costanti concentrate	7
1.3.2	Sistemi acustici distribuiti	8
1.3.3	Il problema agli autovalori	9
1.3.4	Ambienti parallelepipedali ideali	10
1.3.5	Ordine teorico del denominatore comune	12
1.3.6	Conclusioni	12
1.4	Il problema dell'ordine	13
1.4.1	Introduzione	13
1.4.2	Un nuovo metodo per la stima dell'ordine	15
1.4.3	Simulazioni	16
1.5	Conclusioni	22
2	Cancellazione d'eco stereo e CAP	29
2.1	La cancellazione d'eco: background	30
2.1.1	Una semplice applicazione per filtri adattativi	30
2.1.2	Il metodo LMS come riferimento teorico	32
2.1.3	L'uso di CAP nell'AEC monocanale	35
2.2	La cancellazione d'eco stereofonica	39
2.2.1	Descrizione del problema	39
2.2.2	Background analitico	41
2.2.3	Una nuova architettura basata su CAP	49
2.3	Set-up sperimentale e risultati	51
2.4	Conclusioni	55

3	Verso un modello analitico con CAP	59
3.1	Introduzione	59
3.2	Poli Comuni per sistemi 1-D	60
3.3	Analisi delle radici in 1-D	62
3.4	Estensione a 3-D	65
4	Metodi di stima del denominatore comune	70
4.1	Problemi aperti nella stima di poli comuni	70
4.2	Nuove idee intorno alla predizione lineare	73
4.3	Predizione Multicanale Vincolata con Burg	78
4.4	Simulazioni e risultati sperimentali	81
4.4.1	Considerazioni preliminari	81
4.4.2	Descrizione del set-up e delle prove	82
4.4.3	Risultati	83
4.5	Conclusioni	84

Introduzione

Questa tesi raccoglie i più rilevanti risultati del lavoro svolto dall'autore nel quadro del XXI ciclo del Dottorato di Ricerca in Ingegneria dell'Informazione e della Comunicazione nel triennio accademico 2005-2008 presso la Facoltà di Ingegneria della Sapienza di Roma.

L'argomento centrale, e insieme il filo conduttore del materiale presentato, è un metodo numerico di rappresentazione di funzioni di trasferimento acustiche detto a *Poli Acustici Comuni*, e abbreviato nel seguito con l'acronimo CAP (dall'inglese *Common Acoustical Poles*), originariamente introdotto nel 1992 [1]. Tale metodo, già usato con successo in questo dipartimento [2] promette di rappresentare gran parte delle proprietà acustiche di un ambiente confinato per mezzo di una componente ricorsiva comune a tutti i canali acustici entro l'ambiente scelto (ovvero ad ogni accoppiamento lineare tra una sorgente e un sensore).

Una simile prospettiva riveste notevole interesse per tutte quelle applicazioni, in ambiti come le Telecomunicazioni, l'Intrattenimento e la Multimedialità più in generale, che sfruttano interfacce acustiche, mobili o multicanale, per la riproduzione o la cattura di contenuti sonori. Il riferimento va in particolare a quella classe detta di "interfacce acustiche intelligenti" [3], che incorpora complessi sistemi di elaborazione del segnale per compiti quali lo *speech enhancement*, l'inversione acustica, la cancellazione d'eco e di rumore, la separazione di sorgenti o la localizzazione sonora.

Sebbene dal momento della sua introduzione la modellazione basata su CAP non abbia trovato frequentissima applicazione, l'avvicinamento al problema all'inizio di questo lavoro ha suggerito che numerose implicazioni dietro l'idea originale fossero ancora inesplorate. In accordo con tali convincimenti, la presente tesi si articola intorno a quattro principali contributi originali, raggiunti in seguito a una nuova e migliore comprensione del significato della modellazione basata su CAP: l'introduzione di un nuovo metodo per la selezione dell'ordine di modellazione [4], l'uso dei CAP nella cancellazione d'eco stereofonica [5], la derivazione di espressioni semi-analitiche per modelli basati su CAP in sistemi 3-D, e la proposta di un nuovo metodo di stima. Proprio attorno a questi risultati è organizzata la divisione del materiale presentato.

Il capitolo 1 introduce la modellazione basata su CAP e i relativi metodi di stima, ne fornisce una motivazione teorica e, dopo aver trattato il problema della scelta dell'ordine di modellazione in questo contesto, propone un nuovo criterio per la sua selezione. La trattazione è seguita dalla descrizione delle simulazioni usate nella validazione del metodo proposto.

Il capitolo 2 si occupa dell'applicazione della modellazione basata su CAP al problema della cancellazione d'eco stereofonica. Il paragrafo 2.1 e parte del 2.2 forniscono una panoramica sulle particolari problematiche connesse a questa particolare applicazione, nel quadro della più generica cancellazione d'eco adattativa. Una nuova architettura basata su CAP è proposta nel paragrafo 2.2.3, e lì confrontata sperimentalmente con il corrispondente sistema allo stato dell'arte in letteratura.

Il capitolo 3 prende le mosse dall'approfondimento della teoria dietro l'idea dei CAP. A dispetto delle difficoltà ricordate nel capitolo 1 sulla derivazione di espressioni analitiche di funzioni di trasferimento acustiche per sistemi 3-D, il lavoro qui presentato ottiene su questo fronte dei risultati nel caso 1-D, e li estende a una particolare configurazione per ambienti a 3-D. Ciò permette di derivare nuove espressioni analitiche per modelli basati su CAP, da una parte allargando la conoscenza del problema in sé, e dall'altra aprendo nuove potenzialità per la valutazione quantitativa dei metodi di stima esistenti.

Nel capitolo 4 è proposto un nuovo semplice metodo di stima dei parametri invarianti nella modellazione CAP, nell'ottica di rispondere ad esigenze non soddisfatte dai metodi di stima attualmente esistenti. I risultati descritti nel capitolo 3 sono qui usati per confrontare sperimentalmente le prestazioni del metodo proposto con quelle del più accreditato tra i metodi esistenti. Le prove presentate mostrano in primo luogo che la nuova tecnica permette di raggiungere un grado di accuratezza considerevolmente superiore. In secondo luogo, essendo le prove eseguite di per sé innovative, esse hanno permesso di mettere in luce importanti debolezze del metodo usato come riferimento.

Capitolo 1

Modelli di sistemi multicanale basati su Poli Acustici Comuni

1.1 Introduzione

Il paradigma della modellazione a poli comuni è stato originalmente introdotto da Haneda e altri autori in un articolo [6] del 1994. L'idea centrale nasce dall'ipotesi che una funzione di trasferimento acustica, relativa all'accoppiamento tra un sorgente e un sensore in un generico ambiente confinato (nel seguito abbreviata con l'acronimo RTF, dall'inglese *Room Transfer Function*), sia modellabile in modo ragionevolmente accurato in termini ARMA (da *Auto-Regressive and Moving Average*), ovvero secondo la seguente generale espressione nel dominio z :

$$H(z) = \frac{B(z)}{A(z)} = \frac{b_0 + b_1 z^{-1} + \dots + b_{\nu_B} z^{-\nu_B}}{1 + a_1 z^{-1} + \dots + a_{\nu_A} z^{-\nu_A}} \quad (1.1)$$

A partire da tale ipotesi, l'assunzione centrale consiste nell'ammettere che i coefficienti del denominatore (e così anche le sue radici, ovvero i poli) possano essere considerati i medesimi per ogni possibile RTF in tale ambiente confinato (ossia per ogni coppia di arbitrarie posizioni di, rispettivamente, sorgente e sensore). Introducendo quindi l'indice di canale i associato alla singola RTF $H_i(z)$, l'espressione analitica di quest'ultima è data da

$$H_i(z) = \frac{B_i(z)}{A(z)} = \frac{b_0^{(i)} + b_1^{(i)} z^{-1} + \dots + b_{\nu_B}^{(i)} z^{-\nu_B}}{1 + a_1 z^{-1} + \dots + a_{\nu_A} z^{-\nu_A}} \quad (1.2)$$

Mentre i coefficienti del numeratore, $b_k^{(i)}$ sono riconosciuti dipendenti dalla singola scelta del canale, il vettore dei coefficienti $\mathbf{a} = [a_1, \dots, a_{\nu_A}]$ è supposto

costante al variare dell'indice i .

Il primo obiettivo dichiarato di questo approccio modellistico è facilmente individuabile: ammettendo di poter conoscere il denominatore comune mediante una stima preventiva, il numero rimanente di parametri necessari alla descrizione completa di un qualsiasi canale acustico nell'ambiente assegnato sarebbe pari al numero di coefficienti del numeratore; questi ultimi sono potenzialmente in numero assai inferiore ai coefficienti necessari a una descrizione mediante la risposta impulsiva. Rispetto all'uso di strutture non ricorsive, inoltre (cioè con filtro FIR costruito per approssimazione della risposta impulsiva) la modellazione a poli acustici comuni (in inglese *Common Acoustical Poles*, o più semplicemente CAP) offre come altro immediato vantaggio la produzione di risposte impulsive illimitate (ipotesi ben più realistica nella modellazione di canali acustici).

Altra interessante osservazione, forse a carattere più scientifico che tecnologico, risiede nell'indicazione che i coefficienti del denominatore possano racchiudere tutte le informazioni necessarie alla caratterizzazione del comportamento acustico di un ambiente assegnato, e in tal modo costituire una funzione invariante dell'ambiente stesso, in grado di incapsulare parametri e gradi di libertà di quest'ultimo.

Dalle precedenti osservazioni dovrebbe apparire abbastanza chiaro che la strategia modellistica proposta debba rivestire un importante interesse, sia da un punto di vista scientifico-conoscitivo, che da un punto di vista tecnologico. Nel campo delle telecomunicazioni e della multimedialità, sono numerose le applicazioni legate all'uso di algoritmi di stima di canali acustici, sia in termini statici che adattativi. Tipici esempi sono le cancellazioni d'eco e di rumore, l'equalizzazione di canale o la localizzazione acustica. Per tutte queste applicazioni, sia la riduzione del numero di parametri da stimare, che il potenziale aumento di accuratezza nella stima di RTF costituiscono vantaggi di particolare rilevanza. [2, 7, 8, 9, 10, 11, 12].

Chiarita l'idea di base e il suo potenziale interesse, resta da definire in primis quali siano le giustificazioni teoriche di un tale approccio modellistico, e in secondo luogo quale siano le difficoltà nella stima preventiva del denominatore comune, e la relativa accuratezza raggiungibile. Come delineato in precedenza, tale operazione è necessaria per accedere ai vantaggi offerti dalla modellazione basata su CAP, e una conoscenza dei suoi limiti e delle sue potenzialità è di fondamentale importanza nella valutazione dell'opportunità dell'uso degli stessi CAP.

Seppure delineato in [6], il problema della giustificazione teorica merita un'attenzione particolare, e verrà trattato nel paragrafo 1.3. Una panoramica sui metodi esistenti verrà invece affrontata nel prossimo paragrafo.

1.2 La stima dei poli comuni

Il problema della stima dei CAP, oltre che essere affrontato in [6], ha meritato un buon livello di interesse in lavori successivi [13, 14, 15]. Anche se tutti i contributi hanno introdotto elementi di novità e interesse, il primo tra i metodi proposti in [6], detto lì semplicemente “Least Squares”, e ribattezzato qui e nel seguito HMK dalle iniziali dei suoi tre autori, rimane il più diffuso. Le ragioni di ciò verranno riassunte nel paragrafo 1.2.2, insieme a una panoramica su alcuni altri importanti metodi esistenti.

1.2.1 Il metodo HMK

Il metodo HMK parte dall'ammettere la conoscenza di N_H risposte impulsive relative a uno stesso ambiente acustico confinato, supposte tutte, senza perdita di generalità, di lunghezza efficace pari a L_H , e rappresentate in forma vettoriale come

$$\mathbf{h}^{(i)} = \left[h^{(i)}[0], \dots, h^{(i)}[L_H - 1] \right]^t \quad \text{per } i = 1, \dots, N_H \quad (1.3)$$

dove la notazione $[\]^t$ rappresenta l'operatore di trasposizione matriciale. Nel dominio z , gli N_H canali sono modellati da N_H funzioni di trasferimento del tipo

$$H^{(i)}(z) = B^{(i)}(z)/A(z) \quad (1.4)$$

in cui $A(z) = 1 + a_1 z^{-1} + \dots + a_{L_A} z^{-L_A}$ rappresenta il denominatore comune, di ordine assunto pari ad L_A , e $\hat{B}^{(i)}(z) = b_0^{(i)} + \dots + b_{L_B-1}^{(i)} z^{-(L_B-1)}$ il numeratore per l' i -esimo canale, di ordine $L_B - 1$.

L'insieme dei parametri da stimare, in numero $L_A + N_H \times L_B$, è raggruppabile nei seguenti $N_H + 1$ vettori:

$$\begin{aligned} \mathbf{a} &= \left[a_0^{(i)}, \dots, a_{L_A-1}^{(i)} \right] \\ \mathbf{b}^{(i)} &= \left[b_0^{(i)}, \dots, b_{L_B-1}^{(i)} \right] \quad \forall i = 1 \dots N_H \end{aligned} \quad (1.5)$$

Per ogni canale, è costruito un errore dato dalla seguente approssimazione della differenza tra la risposta impulsiva misurata e quella prodotta dal modello:

$$\varepsilon_{(eq)}^{(i)}[n] = h^{(i)}[n] + \sum_{j=1}^{L_A} \hat{a}_j h^{(i)}[n-j] - \sum_{\ell=1}^{L_B} \hat{b}_\ell^{(i)} \delta[n-\ell] \quad (1.6)$$

Si noti come, per mantenere una dipendenza lineare tra l'errore e le $a_k^{(i)}$, la risposta misurata $h^{(i)}[n-j]$ sia usata per esprimere il contributo della componente ricorsiva del modello. Al fine di determinare i parametri, la seguente norma quadratica è costruita per l'errore di modellazione, sommando i contributi di tutti i canali:

$$\varepsilon_2 = \sum_{i=1}^{N_H} \sum_{n=0}^{L_H-1} \left(\varepsilon_{(eq)}^{(i)}[n] \right)^2 \quad (1.7)$$

Le a_k e le $b_j^{(i)}$, sono determinate imponendo che la (1.7) sia minima. Ne risulta un sistema lineare, per la cui soluzione è necessario calcolare la pseudo-inversa una matrice di dimensioni

$$N_H (L_A + L_H + 1) \times [L_A + N_H (L_B + 1)] \quad (1.8)$$

In [6], oltre alla dettagliata espressione del sistema risolutivo, è rintracciabile una parziale prova analitica della stabilità e dell'unicità delle soluzioni così definite. Partendo inoltre dalla conoscenza delle risposte impulsive degli N_H canali, un vantaggio di questo metodo è la relativa semplicità realizzativa.

D'altro canto tale semplicità, derivante dal consistere unicamente nella soluzione di un sistema lineare di forma nota, finisce per essere anche l'origine dei principali punti negativi. In particolare, le elevate dimensioni della matrice da invertire lo rendono poco pratico anche al crescere di complessità estranee al modello in sé, come il numero di canali, N_H e la lunghezza delle risposte impulsive misurate disponibili, L_H .

Alcune alternative disponibili in letteratura sono presentate nel paragrafo successivo. Il capitolo 4 introdurrà invece un nuovo metodo, che costituisce uno dei contributi originali di questo lavoro. E' comunque importante notare che il metodo HMK è comunemente individuato in letteratura come il metodo di riferimento e, per complessità modellistiche contenute, possiede le necessarie caratteristiche di robustezza e semplicità per essere usato senza troppe accortezze. Il capitolo 4 presenterà anche uno studio quantitativo sulle proprietà di accuratezza del metodo HMK, probabilmente non esplorate fino in fondo in [6] né in altri lavori.

1.2.2 Altri metodi per la stima di poli comuni

Proprio a causa della potenziale insostenibilità computazionale, gli stessi autori del metodo HMK propongono (e usano) in [6] un metodo alternativo, da adottare quando l'ordine dei parametri in gioco non consente di usare il primo. Tale approccio consiste in pratica nella determinazione di N_H semplici stime ARMA sugli N_H canali disponibili (sempre in termini di risposte impulsive). Usando una notazione simile a quella che definita dalle (1.5), ciò porta a determinare N_H coppie di vettori $\mathbf{a}^{(i)}$, $\mathbf{b}^{(i)}$, con $i = 1 \dots N_H$, da cui la stima finale del vettore dei coefficienti del denominatore comune \mathbf{a} è ottenuta come semplice media delle stime AR di singolo canale $\mathbf{a}^{(i)}$. Sempre in [6] l'accuratezza di questo metodo "ridotto", battezzato lì *Averaging Method* (nel seguito AM) per ovvie ragioni, è riconosciuta ampiamente inferiore al metodo HMK.

Zuccarello e altri propongono in [15] una strategia di calcolo dei CAP che prende le mosse, come l'AM, dalle stime ARMA di singolo canale. Rispetto al primo, il metodo lì proposto procede però prima all'estrazione delle radici del polinomio AR stimato, e successivamente identifica i poli comuni mediante una

ricerca di *cluster* per i poli ricavati dall'insieme dei canali. Nonostante l'interessante contributo, quest'ultimo metodo trova scarso riscontro in letteratura. Le motivazioni sono in primo luogo da ricercarsi nell'eccessivo incremento di complessità computazionale (introdotto dalla necessità di estrazione delle radici), e in secondo luogo della necessità di usare di algoritmi di ricerca di *cluster*. Per questi ultimi, d'altro canto, rimane non chiarita l'effettiva capacità di discriminazione all'aumento dell'ordine e, come sarà più chiaro in seguito, della densità dei poli stessi nel piano z .

Un ultimo importante contributo al problema della stima dei CAP è proposto da Kikichi e Miyoshi nel 2004 [13, 14]. In questi interessanti lavori, la ricerca del denominatore comune $A(z)$ è associata a quella del polinomio caratteristico di una “matrice di predizione lineare”. Quest'ultima è costruita a partire dai segnali acquisiti come uscite di un sistema acustico SIMO (dall'inglese *Single Input Multiple Output*) all'interno dell'ambiente considerato. Oltre all'importante derivazione teorica, questo metodo porta con sé il grande vantaggio di non necessitare della stima preventiva delle N_H risposte impulsive. Il non richiedere formalmente la conoscenza del segnale usato come ingresso del sistema SIMO fa di questo approccio un metodo *Blind* (o anche, in italiano, “non supervisionato”). Nonostante questi punti a favore, lo spazio per l'utilizzo pratico di questo metodo è assai limitato: l'insieme di problemi di stabilità e impraticabilità computazionali lo rendono inutilizzabile sia per ordini elevati che per un numero di canali superiore a due.

La panoramica offerta in questo paragrafo è intesa a dare un'idea sul perché l'HMK rimanga a tutt'oggi il metodo più usato nella stima di CAP. In secondo luogo, gli spunti di interesse associati ad altri metodi esistenti, e gli svantaggi che l'HMK continua ad avere, fornisce una prima spiegazione dell'esigenza di ricercare nuove strategie per la stima di CAP.

1.3 La giustificazione teorica dei modelli basati su CAP

1.3.1 Denominatore comune e reti a costanti concentrate

Il lettore familiare con la teoria delle reti a costanti concentrate (RCC) non ha probabilmente trovato particolarmente inusuale l'approccio modellistico ARMA basato su CAP, come introdotto all'inizio di questo capitolo. Per un generico sistema a costanti concentrate sono definibili un certo numero di *funzioni di rete* $H^{(i)}$ attraverso il rapporto tra grandezze (per reti elettriche tensioni o correnti, rispettivamente) in un dominio trasformato e definite in due punti del sistema in generale diversi. Il concetto di funzione di rete in sé è del tutto analogo a quello di funzione di trasferimento adottato in campo acustico, se le due

grandezze considerate sono i due segnali di pressione rispettivamente emesso da una sorgente e ricevuto da un sensore.

Un primo fondamentale contributo dato dalla teoria delle RCC [16, 17] è che, nel dominio della trasformata di Laplace (qui e nel seguito ci si riferirà alla frequenza complessa di Laplace con il simbolo s), qualsiasi funzione di rete ha forma razionale fratta, cioè è esprimibile come rapporto tra due polinomi a coefficienti costanti nella variabile s :

$$H'(s) = \frac{B'(s)}{A'(s)} = \frac{\beta_0 + \beta_1 \cdot s + \beta_2 \cdot s^2 \dots}{1 + \alpha_1 \cdot s + \alpha_2 \cdot s^2 \dots} \quad (1.9)$$

La seconda e forse ancora più importante informazione proveniente sempre dalla teoria delle RCC è che qualsiasi funzione di rete relativa al medesimo sistema condividerà lo stesso denominatore $A'(s)$: è qui il nodo che sembra giustificare la modellazione basata su CAP. Estendendo le considerazioni basate sulle RCC ai sistemi acustici, e rispettando il parallelo tra funzioni di rete e RTF, l'uso dei CAP potrebbe apparire ragionevolmente giustificato. In primo luogo, infatti, l'uso di un mapping lineare o bilineare (come trasformazione bilineare $s = (2/T_C) \cdot (1 - z^{-1})/(1 + z^{-1})$) nel passaggio dal dominio s a quello della trasformata z conserva la forma analitica razionale, nonché l'indipendenza tra poli e zeri: se i poli sono comuni in s , tali rimarranno in z . Se dunque le RTF sono stimate in forma ARMA, cioè come precisamente funzioni razionali fratte in z , sembra naturale imporre al denominatore di non subire alcun cambiamento per diverse RTF relative a uno stesso ambiente acustico.

Prima di spiegare perché tale assunzione non sia realmente lecita in questa semplice forma, sembra utile ricordare un'altra specificità del denominatore delle funzioni di rete di RCC. Gli α_i definiti della (1.9) sono sempre:

- Funzioni razionali dei parametri concentrati del sistema;
- In numero pari esattamente al numero di elementi reattivi (o con memoria) indipendenti presenti nel sistema.

Varrà quindi innanzitutto sempre

$$\begin{cases} b_i = f_i(\alpha, \beta) = b_i(\{R\}, \{L\}, \{C\}) \\ a_j = g_j(\alpha, \beta) = a_j(\{R\}, \{L\}, \{C\}) \end{cases} \quad (1.10)$$

e inoltre la descrizione del sistema in termini del suo *polinomio caratteristico* $A(z)$ sarà sempre una descrizione massimamente *parsimoniosa*, perché in grado di essere definita da un numero di coefficienti esattamente pari al numero di gradi di libertà del sistema fisico di riferimento.

1.3.2 Sistemi acustici distribuiti

Le considerazioni fatte fin qui non sono in realtà sufficienti a giustificare la modellazione basata su CAP per sistemi acustici compatibili con la presenza umana. Il

nodo fondamentale della questione è nelle dimensioni spaziali dei sistemi normalmente considerati come ambienti acustici. Un sistema è lecitamente studiabile usando un approccio a costanti concentrate quando, della ℓ la sua massima dimensione lineare, t_{\min} la risoluzione temporale e c la velocità di propagazione del fenomeno ondulatorio d'interesse (in questo caso la velocità del suono in aria), è soddisfatta la seguente disuguaglianza [16]

$$c \cdot t_{\min} \gg \ell \quad (1.11)$$

o anche, usando $t_{\min} = 1/(2f_B)$, con f_B la più grande frequenza d'interesse (o larghezza di banda, o frequenza di Niquist con esplicito riferimento al campionamento)

$$2f_B \frac{\ell}{c} \ll 1 \quad (1.12)$$

Ipotizzando ℓ non inferiore a qualche unità di metri, $c \simeq 340 \text{ ms}^{-1}$ ed f_B pari anche solo a qualche kHz, si può facilmente verificare che la (1.12) non è mai verificata. In altri termini, i sistemi acustici ambientali e i fenomeni correlati non sono mai spiegabili con la teoria delle RCC a causa della loro natura intrinsecamente *propagativa*: le dimensioni spaziali sono fondamentali, accanto a quella temporale, per la descrizione delle grandezze d'interesse (pressione e corrente) e per la definizione dei parametri che definiscono il sistema stesso.

Da un punto di vista analitico, la descrizione non è basata su equazioni differenziali ordinarie (o ODE dall'inglese Ordinary Differential Equation), ma da equazioni alle derivate parziali (PDE). Nello specifico, la PDE coinvolta è l'equazione delle onde, che in condizioni stazionarie e nel dominio di Laplace è riconducibile all'equazione di Helmholtz, riportata qui di seguito nella sua forma non omogenea [18]

$$\nabla^2 p(s; \mathbf{x}) - s^2 p(s; \mathbf{x}) = -s\rho_0 q(\mathbf{x}) \quad (1.13)$$

in cui $p(s; \mathbf{x})$ rappresenta la trasformata della pressione osservata in un punto \mathbf{x} , in presenza di una sorgente di densità spaziale $q(\mathbf{x})$. La (1.13) consente di definire agevolmente una formulazione analitica di RTF nel dominio s tramite

$$H(s) = \frac{p(s; \mathbf{x})}{q(\mathbf{x})} \quad (1.14)$$

1.3.3 Il problema agli autovalori

È importante osservare che la (1.13) e la (1.14) non garantiscono più una forma razionale fratta per la $H(s)$: le soluzioni avranno anzi forma generalmente trascendente in s . Una relativa semplificazione del problema emerge dopo l'imposizione delle condizioni al contorno e la definizione delle autofunzioni spaziali della (1.13), $H_k(\mathbf{x})$ per $k = 1 \dots \infty$ [18]: detta \mathbf{x}_s la posizione della sorgente e \mathbf{x}_r quella del sensore, lo sviluppo in serie di $q(\mathbf{x})$ lungo le $H_k(\mathbf{x})$ consente di

definire una soluzione per la RTF $H(s; \mathbf{x}_s, \mathbf{x}_r)$ a sua volta tramite il seguente sviluppo

$$H_L(s; \mathbf{x}_s, \mathbf{x}_r) = s \cdot d \cdot \sum_{k=1}^{\infty} H_k(\mathbf{x}_s) H_k(\mathbf{x}_r) \left(\frac{s^2}{\Omega_k^2} + \frac{s}{\Omega_k Q_k} + 1 \right)^{-1} \quad (1.15)$$

Nella precedente espressione le Ω_k sono le pulsazioni modali associate ad $H_k(\mathbf{x})$, e Q_k il corrispondente fattore di merito, o di qualità.

La (1.15) ha il grande merito di riaprire uno spiraglio sulla giustificazione dell'uso dei CAP. Se la sommatoria si limitasse a visitare un numero finito di termini, i singoli denominatori rappresenterebbero i singoli fattori del denominatore comune, non contenendo nessuna dipendenza funzionale né da \mathbf{x}_s né da \mathbf{x}_r , ossia dalle rispettive posizioni di sorgente e ricevitore, che identificano il particolare canale.

In effetti, come verrà chiarito tra breve, i termini razionali del secondo ordine della (1.15) corrispondono a valori di Ω_k via via crescenti all'aumentare di k ; in particolare il contributo di ognuno di essi non è particolarmente significativo per $f \ll \Omega_k/(2\pi)$ o per $f \gg \Omega_k/(2\pi)$. E' possibile pertanto ipotizzare che, se lo spettro dei segnali coinvolti è contenuto entro una banda f_B , i termini del secondo ordine corrispondenti a $\Omega_k > \Omega_B = 2\pi f_B$ non debbano contribuire in modo decisivo a $H_L(s)$. In tal modo è possibile affermare che la rappresentazione di RTF con modelli ARMA (e in particolare a CAP) è giustificata in termini approssimati, e per problemi a banda limitata. Quest'ultima ipotesi è peraltro più che naturale, visto che la nozione stessa di modellazione ARMA ha senso esclusivamente nel tempo discreto: ciò permette l'assunzione automatica $f_B = F_C/2$, essendo F_C la frequenza di campionamento usata.

1.3.4 Ambienti parallelepipedali ideali

La definizione di una particolare geometria consente una comprensione più accurata di quali siano le espressioni analitiche delle pulsazioni modali Ω_k , e di conseguenza quali siano gli ordini tipici coinvolti nel troncamento dello sviluppo (1.15). Risultati comunemente noti, e peraltro assai pertinenti per il tipo di applicazioni obiettivo dell'uso di CAP, sono quelli associati all'ipotesi di ambiente confinato a forma di parallelepipedo. In particolare, la soluzione del problema degli autovalori è facilmente accessibile in ipotesi di ambiente completamente non dissipativo, e quest'ultima si assume ragionevolmente estendibile anche in presenza di bassa dissipazione alle pareti [18].

Dette L_1 , L_2 ed L_3 le tre rispettive dimensioni della stanza, le frequenze modali sono date dalla seguente espressione

$$F_{k_1, k_2, k_3} = \frac{c}{2} \sqrt{\left(\frac{k_1}{L_1}\right)^2 + \left(\frac{k_2}{L_2}\right)^2 + \left(\frac{k_3}{L_3}\right)^2} \quad k_i = 0, 1, \dots, \infty \quad (1.16)$$

in cui c è come al solito la velocità del suono, e le triplette (k_1, k_2, k_3) sono definite da tutte le possibili combinazioni di interi non negativi. Fissate le dimensioni della stanza, è possibile contare in numero di triplette (k_1, k_2, k_3) che definiscono frequenze modali comprese entro f_B . Alternativamente esistono espressioni approssimate, tanto più accurate quanto più grandi le k_i . A tal proposito un'attenzione particolare merita la nota formula [18]

$$n_M(f_B; L, S, V) \simeq \frac{4}{3}\pi V \left(\frac{f_B}{c}\right)^3 + \frac{\pi}{4}S \left(\frac{f_B}{c}\right)^2 + \frac{L}{8} \left(\frac{f_B}{c}\right) \quad (1.17)$$

che esprime in modo approssimato il numero di frequenze modali n_M comprese entro una frequenza di banda f_B , per una stanza parallelepipedale avente somma delle lunghezze degli spigoli pari a L , superficie complessiva S e volume V . Alternativamente, la (1.17) può essere espressa attraverso la sua derivata rispetto a f_B , e definire così la *densità modale* sull'asse della frequenza

$$\delta_M(f) \triangleq \frac{dn_M(f)}{df} = 4\pi V \frac{f^2}{c^3} + o(f) \quad (1.18)$$

Un'utile rappresentazione grafica della (1.17) è presentata in figura 1.1 per bassi valori di f_B . L'osservazione della distribuzione delle frequenze modali sull'asse delle frequenze fornisce inoltre una significativa interpretazione visiva della (1.18), data dalla densità delle linee verticali.

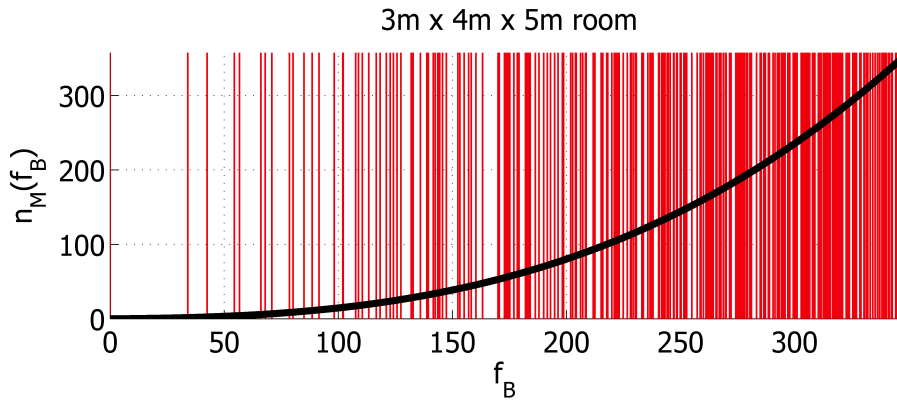


Figura 1.1: Andamento del numero di modi di vibrazione all'aumentare della frequenza di banda f_B per una stanza di dimensioni $3 \times 4 \times 5 \text{ m}^3$. La curva in nero è data dalla (1.17), mentre le righe verticali rappresentano tutte le frequenze modali individuate dalla (1.16).

1.3.5 Un ordine teorico per il denominatore comune in sistemi acustici

Gli ordini di grandezza generati dalle (1.16), (1.17) e (1.18) per le dimensioni lineari usate in figura 1.1 danno facilmente nell'occhio. Righe spettrali ad alta densità sono in generale difficili da stimare e discriminare, come suggerisce la teoria dei comuni metodi parametrici di stima spettrale [19]: come regola generale, la capacità di definizione di righe spettrali vicine tra loro è progressivamente compromessa all'aumentare della densità delle righe stesse. In tal quadro è difficile pensare che metodi come l'HMK descritto in 1.2.1 possano fare significative eccezioni rispetto alla stima di CAP, quando la densità modale (o similmente la densità nel piano z delle coppie di poli complessi coniugati) cresce come in $O(f^2)$. È bene infatti ricordare che l'ordine CAP prescritto dalla teoria (pur in ipotesi di descrizione approssimata per troncamento della (1.15)) è dato da

$$L_A = 2 \times n_M(f_B; L, S, V) \quad (1.19)$$

In aggiunta al problema della densità, la panoramica sui metodi di stima dei CAP, e in particolare l'introduzione al metodo HMK fatta nel paragrafo 1.2.1, aggiunge un altro particolare degno di nota: il metodo HMK si basa sull'inversione di una matrice di dimensioni date dalla (1.8), qui riportata per comodità

$$N_H (L_A + L_H + 1) \times [L_A + N_H (L_B + 1)]$$

Entrambe le dimensioni di tale matrice aumentano linearmente con L_A , e quindi secondo f_B^3 . Ne deriva un altro pesante indizio sulla grande difficoltà di stima dei CAP in sistemi e dimensionalità verosimili.

1.3.6 Conclusioni

Come verrà spiegato nel paragrafo 1.4, più che rendere proibitivo l'uso dei CAP, il problema dell'adottabilità dell'ordine teorico apre una riflessione su quale debba essere la scelta di un ordine adottabile in pratica, che consenta, da una parte, di individuare il maggior numero possibile di coppie di poli complessi coniugati descriventi le proprietà risonanti del sistema acustico, e dall'altra, di usare metodi di stima come l'HMK disponibili in letteratura.

Oltre ai problemi sulla scelta di un ordine pratico, le considerazioni fatte fin qui sulla stima dei CAP lasciano ancora aperti numerosi interrogativi. Schematicamente, il seguente elenco può fornirne un primo raggruppamento.

- **Incidenza e ragionevolezza dell'approssimazione ad ordine finito:** l'assunzione che il troncamento della (1.15) produca risultati soddisfacenti per problemi a banda limitata non è stata giustificata in modo sufficientemente dettagliato. L'assenza di una soluzione analitica per tutte

le soluzioni dell'equazione di Helmholtz in problemi tridimensionali rende difficile capire in che termini possa assumersi lecita la descrizione razionale fratta.

- **Nessun ordine teorico per i numeratori del modello:** la (1.15) suggerisce che i numeratori della descrizione razionale fratta approssimata debbano avere un ordine pari a quello del denominatore comune, ma la stessa formula non ne suggerisce alcuna particolare interpretazione fisica; se le radici del denominatore comune hanno un'interpretazione sufficientemente intuitiva, è difficile poter dire altrettanto dei numeratori di singolo canale o delle loro radici.
- **Parsimoniosità descrittiva:** in contrapposizione a quanto si è detto per la modellazione a CAP di funzioni di rete per sistemi a costanti concentrate, la rappresentazione CAP di sistemi acustici sembra perdere la tanto desiderabile proprietà della massima parsimoniosità. In altri termini, sebbene un sistema acustico possa essere potenzialmente descritto da un numero assai limitato di parametri (ad esempio le tre dimensioni lineari e i coefficienti - potenzialmente costanti - descriventi le proprietà dissipative delle pareti), la sua descrizione (e stima) mediante CAP sembra difficilmente riconducibile alla stima di un numero ridotto di parametri, e appare quindi assai scarsamente parsimoniosa.

Vale infine la pena ricordare che gli approfondimenti fatti fin qui, sebbene in parte già suggeriti in [6], non sono stati mai riassunti in passato in modo unitario e sistematico.

1.4 Il problema dell'ordine

1.4.1 Introduzione

Ogni stima dei parametri di un modello a più canali basato su CAP, come quello individuato dalle (1.2), necessita in primo luogo di un'ipotesi sugli ordini del modello da stimare. Nel caso in questione, gli ordini sono definiti da una scelta per la coppia L_A, L_B pari, rispettivamente all'ordine del denominatore dei poli comuni, e alla lunghezza delle componenti MA di ogni singolo canale. Se un aspetto di cruciale importanza legato alla scelta dell'ordine deriva dal definire la complessità computazionale del problema di stima (come definito dalla (1.8)), la sua vera importanza è associata a metriche quali:

1. Il livello di parsimoniosità
2. Il grado di accuratezza descrittiva
3. La corrispondenza analitica tra il modello e il sistema fisico reale

Mentre i due primi criteri trovano applicazione in ogni ambito modellistico, anche di carattere puramente comportamentale (cioè nei casi in cui il modello non ha una struttura matematica direttamente ispirata alla fisica del sistema reale a cui si riferisce), il terzo è proprio dei modelli cosiddetti “strutturali”^S, cioè per i quali esiste una corrispondenza diretta tra la formalizzazione del modello e la descrizione analitica del fenomeno fisico descritto. Per i modelli strutturali è naturalmente buona norma, ove possibile, attribuire al modello stesso un numero di parametri pari al numero di gradi di libertà che descrivono il sistema fisico modellato.

Potendo puntare a privilegiare il terzo criterio, si raggiunge autonomamente un buon equilibrio nei primi due, normalmente in reciproca contrapposizione. Viceversa, quando la corrispondenza con il modello fisico non può fornire indicazioni risolutive, è proprio l’equilibrio tra la necessità di massimizzare la parsimoniosità e l’esigenza di raggiungere un’elevata accuratezza descrittiva dei fenomeni che costituisce l’ingrediente principale dei tradizionali criteri di scelta dell’ordine, come l’AIC di Akaike [20].

Nel caso dei CAP la componente strutturale (la corrispondenza con il modello fisico) riveste un’importanza di primo piano. Nonostante ciò, perché la coppia L_A, L_B rifletta la complessità del sistema fisico reale, i valori dei due ordini dovrebbero essere così alti da generare due tipi di problemi:

1. Il metodo HMK sarebbe computazionalmente troppo oneroso a causa delle dimensioni descritte dalla (1.8), e in maniera simile probabilmente ogni altro metodo di stima;
2. Anche riuscendo a gestire la complessità computazionale, il metodo di stima sarebbe difficilmente in grado di restituire risultati accurati, e anzi si esporrebbe facilmente al fenomeno dell’*overmodeling* [21]: una parte dei parametri stimati (qui di fatto poli del denominatore comune) non sarebbe significativa.

Riportato al caso specifico della stima dei CAP e del loro successivo uso, il fenomeno dell’*overmodeling* è particolarmente indesiderabile: se, disponendo di N_H risposte impulsive campione, si stimano un numero di CAP superiore a quelli effettivamente stimabili (una sorta di ordine efficace, derivante dalla definizione contemporanea di sistema fisico e metodo di stima), usando il denominatore comune per la stima dinamica di *altre* RTF si rischia di non ottenere affatto un effetto di diminuzione di parametri liberi, perché alcuni dei nuovi zeri dinamicamente stimati per la parte MA andranno a compensare i poli del denominatore comune erroneamente stimati.

Per capire meglio le strategie tradizionalmente usate per selezionare un ordine efficace, si definisca il seguente parametro di accuratezza, detto del *disallineamento* tra le risposte impulsive misurate e quelle date dal modello multicanale

a CAP valutabile dopo la stima dei parametri $\hat{\mathbf{a}}, \hat{\mathbf{b}}^{(i)}$ ($i = 1 \dots N_H$) in (1.5):

$$\mu = \frac{\sum_{i=1}^{N_H} \sum_{n=0}^{L_H-1} \|h_i[n] - h_i[n; \hat{\mathbf{a}}, \hat{\mathbf{b}}_i]\|_2^2}{\sum_{i=1}^{N_H} \sum_{n=0}^{L_H-1} \|h_i[n]\|_2^2} \quad (1.20)$$

Nella più semplice delle strategie (presentata ad esempio in [6] insieme al metodo HMK) si definisce a priori un obiettivo di accuratezza $\bar{\mu}$ per μ , e si aumenta l'ordine del modello fino a raggiungere un valore appena inferiore a $\bar{\mu}$; l'ordine così trovato rappresenta l'ordine efficace usato per la stima definitiva dei parametri del modello.

Sfortunatamente, anche se non veramente osservato in [6] né, a conoscenza dell'autore di questo lavoro, in nessun altro contributo presente in letteratura, una scelta guidata da questo criterio si scontra con il problema dell'ambiguità. Il modello in questione è definito dalla coppia di ordini L_A, L_B , che contribuiscono indipendentemente alla definizione del numero di parametri del sistema, ovvero $N_H \times L_B + L_A$. A causa del diverso significato di L_A ed L_B , prove informali hanno mostrato che è possibile sceglierne diverse combinazioni, tutte in grado di portare a risultati praticamente indistinguibili per μ . In questo quadro il metodo AIC (peraltro già scartato in [6]), che bilancia in modo giustificato accuratezza e numero di parametri (parsimoniosità) non aiuta in modo significativo a sciogliere il nodo dell'ambiguità.

1.4.2 Un nuovo metodo per la stima dell'ordine

L'aver introdotto un nuovo criterio per la scelta dell'ordine con l'uso del metodo HMK è stato il primo contributo originale [4] del presente lavoro. L'obiettivo centrale è costituito dalla rimozione dell'ambiguità nella selezione degli ordini L_A ed L_B , possibilmente introducendo nella metrica di accuratezza una componente che tenga conto della specificità del problema trattato: la stima di una componente *comune* ai diversi N_H canali modellati.

Al posto del disallineamento globale in (1.5), si considerino i seguenti disallineamenti di singolo canale

$$\mu_i = \frac{\sum_{n=0}^{L_H-1} \|h_i[n] - h_i[n; \hat{\mathbf{a}}, \hat{\mathbf{b}}_i]\|_2^2}{\sum_{n=0}^{L_H-1} \|h_i[n]\|_2^2} \quad \forall i = 1 \dots N_H \quad (1.21)$$

In questo caso per ogni processo di stima la (1.21) definisce N_H diversi disallineamenti di canale, definiti dalla . Appare ragionevole ipotizzare che indicazioni utili possano derivare dall'osservazione statistica dell'*ensemble* di tali μ_i . In particolare, uno dei più semplici comportamenti osservabili riguarda il *massimo* tra

gli N_H valori μ_i , cioè quello corrispondente al canale che il modello globale stima con la *peggiore* accuratezza. Partendo dall'assunzione che il denominatore stimato, di ordine L_A , debba individuare poli *comuni* a tutti i canali, per un valore fissato di L_B è ragionevole la seguente:

Assunzione \bar{L}_A è una buona stima dell'ordine efficace del denominatore **a** se, al crescere di L_A e ripetendo ogni volta la stima di tutti i parametri in (1.5), esiste almeno un canale i , corrispondente all'accuratezza peggiore sull'ensemble, che non esibisce più alcun miglioramento per $L_A > \bar{L}_A$

Sfruttare quest'assunzione corrisponde a definire un criterio per la scelta di un valore efficace per il denominatore dei poli comuni $A(z)$. Resta ancora indefinita la scelta di L_B . Per quest'ultimo, nella più semplice delle ipotesi è possibile stabilire un obiettivo (soglia minima) per l'accuratezza di modellazione (un valore massimo entro cui debbano rientrare tutti i disallineamenti (1.21)), e scegliere come ordine efficace \bar{L}_B dei numeratori il primo valore per cui questa soglia sia superata da tutto l'ensemble dei μ_i .

Sulla scorta delle osservazioni appena fatte, i seguenti passi riassumono il nuovo metodo introdotto per la stima degli ordini di modellazione. Resta da aggiungere che, se è noto l'ambiente acustico entro cui è definito il sistema multicanale oggetto della stima, esiste sempre un valore *teorico* (descritto dalla fisica del sistema reale) dell'ordine L_A , sempre molto maggiore dell'ordine efficace fornito dal metodo qui presentato, nel senso più volte descritto. Per rendere non ambiguo il criterio presentato, si assumerà disponibile una conoscenza grossolana di quest'ordine teorico, L_A^T .

1. Si fissi per L_A un valore iniziale molto grande, ad esempio $L_A^T/2$, e un obiettivo minimo di accuratezza, $\bar{\mu}$
2. Partendo dal basso, si aumenti L_B e si ripeta ogni volta una stima del modello, fino al primo valore, \bar{L}_B per cui $\mu_i < \bar{\mu} \quad \forall i = 1 \dots N_H$
3. Fissato $L_B = \bar{L}_B$ si vari dal basso L_A fino a quel valore \bar{L}_A tale che, per $L_A > \bar{L}_A$, il massimo dei disallineamenti $\mu^{(i)}$ non diminuisca più rispetto al valore per $L_A = \bar{L}_A$
4. Con $L_A = \bar{L}_A$, si vari nuovamente L_B fino a raggiungere un nuovo valore \bar{L}_B , minore per cui valga $\mu_i < \bar{\mu} \quad \forall i = 1 \dots N_H$

1.4.3 Simulazioni

Delle prove sperimentali sono state compiute per testare la non ambiguità del metodo proposto, e per quantificare quale fosse l'effettivo rapporto tra l'ordine efficace così selezionato e il vero ordine dei poli comuni del sistema fisico di

riferimento. Le valutazioni su quale fosse il tipo di simulazione più adatto alla verifica del metodo proposto hanno puntato a soddisfare il compromesso tra:

- La scelta di canali il più simili possibile a sistemi reali, in modo sollecitare il metodo con test di naturale complessità;
- La scelta di un sistema conoscibile in modo sufficientemente dettagliato da poter effettuare un confronto quantitativo tra i risultati ottenuti e le proprietà del sistema di riferimento.

Per soddisfare il secondo punto si è preferito un contesto simulato; allo stesso tempo, per produrre delle risposte impulsive proprie di un sistema strettamente distribuito, ci si è tenuti alla larga da processi di generazione ricorsivi e si è scelto il metodo delle sorgenti-immagine. In questo modo la simulazione dei fenomeni acustici è a carattere esclusivamente propagativo, e in quanto tale, per le considerazioni fatte nel paragrafo 1.3.2, ideale per provare la praticabilità di una modellazione ARMA, tanto più se basata su CAP.

Set-up sperimentale

Per generare le risposte impulsive con il metodo delle sorgenti-immagine è stato scelto un popolare programma [22] di simulazione acustica sviluppato in ambiente MATLAB. Strumenti di questo tipo permettono di definire dimensioni e proprietà acustiche di un ambiente confinato di forma geometrica sufficientemente semplice (è normale aspettarsi che, come in questo caso, ci si limiti alla forma parallelepipedale), numero e posizioni di sorgenti e sensori, e di avviare, a partire da tali informazioni, un processo di simulazione mirato a calcolare tutte le risposte impulsive definite dagli accoppiamenti introdotti.

Nel caso in questione è stata scelta una stanza rettangolare di $(2 \times 2 \times 3) \text{ m}^3$. Come delineato nella proiezione di figura 1.2, quattro sorgenti sono state posizionate in $y = 2 \text{ m}$, $z = 1.5 \text{ m}$, equispaziate lungo l'asse x , e quattro microfoni in $y = 0.5 \text{ m}$, $z = 1 \text{ m}$ con x rispettivamente pari a $[0, 0.13, 0.25, 0.67, 1.125] \text{ m}$. In totale, il sistema così costruito ha generato 16 risposte impulsive campione. Rispetto alle proprietà assorbenti, in virtù delle libertà lasciate disponibili dal contesto di simulazione, sono stati scelti dei coefficienti di assorbimento tali da rendere le risposte impulsive generate prodotte compatibili con il metodo HMK; ci si riferisce qui in particolar modo alla lunghezza delle risposte impulsive stesse, per l'impatto che tale parametro ha sulle dimensioni (1.8). Ciò ha portato ad attribuire alle pareti verticali un coefficiente di assorbimento in energia costante e pari a 0.9, mentre a pavimento e soffitto sono stati lasciati coefficienti di assorbimento più realistici, variabili in funzione della frequenza.

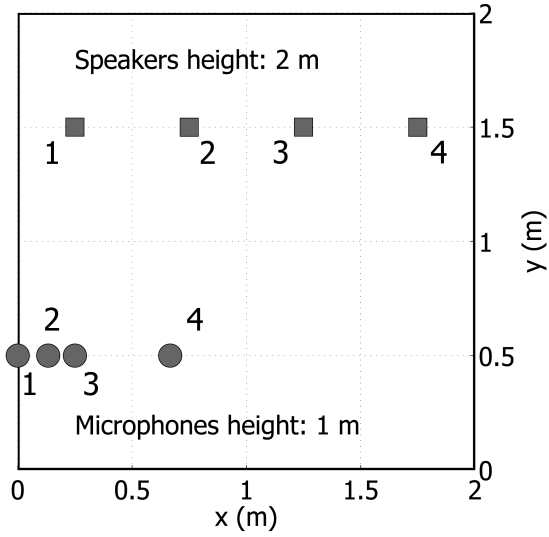


Figura 1.2: Il set-up usato per le simulazioni: quattro sorgenti omni-direzionali (quadrati) e cinque sensori (cerchi), rispettivamente in $x = (0, 0.13, 0.25, 0.67, 1.125)$ m.

Risultati

Tutte le risposte impulsive generate, aventi andamenti come quello in figura 1.3 hanno mostrato un tempo di riverbero tipico $T_{60} \approx 100$ ms.

Le RI $h^{(i)}[n]$ sono state ottenute a una frequenza di campionamento $F_C = 8$ kHz; tutte sono poi state troncate a una stessa lunghezza \bar{n} in grado di soddisfare $\bar{n} \gg T_{60}F_C \forall i = 1, \dots, N_H$. Successivamente, le RI ottenute sono state decimate per i cinque diversi fattori 4, 6, 8, 12, 16, permettendo così di reitarrare il metodo di scelta dell'ordine per cinque diverse larghezze di banda, ossia, rispettivamente $f_B = 1000, 667, 500, 333$ e 250 Hz; ciò al fine di accedere a un andamento sperimentale dell'ordine efficace L_A in funzione di f_B .

Ottenute le 16 risposte impulsive per i diversi valori di f_B , il passo successivo è consistito nell'uso del metodo descritto nel paragrafo 1.4.2, che verrà qui ripercorso per facilitarne una migliore comprensione. L'avere imposto costruttivamente le caratteristiche geometriche dell'ambiente simulato permette di conoscere il valore teorico L_A^T e di usarlo, come indicazinoe, per stabilire un valore di partenza nella ricerca del valore efficace. Fissato L_A , il metodo HMK è applicato iterativamente per diversi valori di L_B fino ad ottenere una distribuzione dei disallineamenti μ_i in (1.21) tutti al di sotto della una soglia di accuratezza minima (o, in altri termini, *disallineamento massimo*) $\bar{\mu}$, in questo caso fissata arbitrariamente a -40 dB.

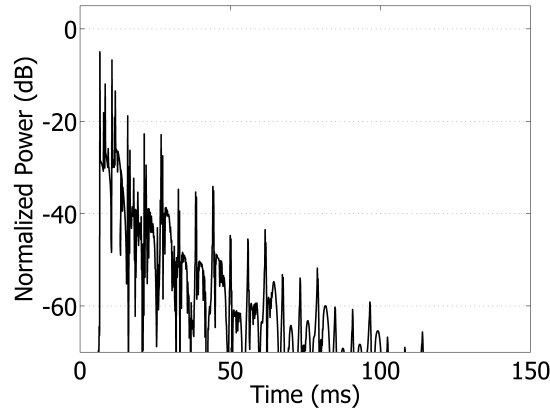


Figura 1.3: Una tipica risposta impulsiva ottenuta del sistema simulato.

Un esempio di come possa apparire l'intervallo di distribuzione dei μ_i al variare di L_B , fino al raggiungimento della soglia, e oltre, è rappresentato in figura 1.4.

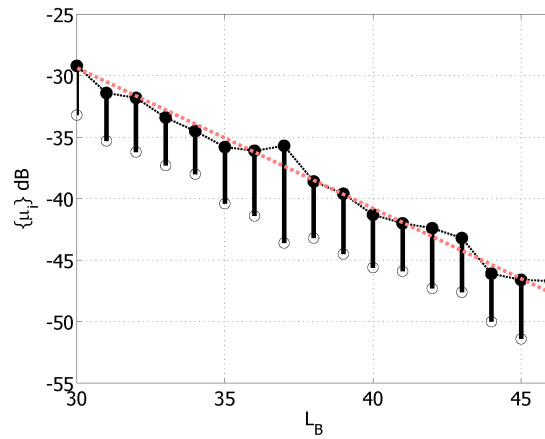


Figura 1.4: Andamento dell'intervallo di distribuzione dei disallineamenti di canale per diverse stime del modello, al variare di L_B , e per $f_B = 500$ Hz.

Fissato $L_B = \bar{L}_B$, pari al minor valore sufficiente ad avere tutti i μ_i entro $\bar{\mu}$, una simile serie di stime è ripetuta al variare di L_A (ad esempio per valori decrescenti, a partire dal primo valore scelto). Come in precedenza, l'intervallo sperimentale entro cui cadono i μ_i è riportato su un grafico, stavolta in funzione di L_A . Ora però l'obiettivo non è più soddisfare un vincolo di accuratezza assoluta, bensì determinare quel valore di L_A che soddisfi il criterio enunciato al punto 3 del paragrafo 1.4.2.

Come può osservarsi in figura 1.5, l'andamento dell'intervallo di distribuzione dei disallineamenti di canale è diverso rispetto al caso in cui si varia L_B . In

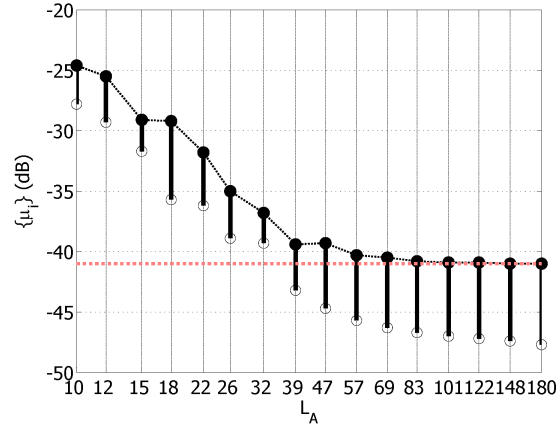


Figura 1.5: Andamento dell'intervallo di distribuzione dei disallineamenti di canale, fissato $L_B = \bar{L}_B$, per diverse stime del modello, al variare di L_A , e per $f_B = 333Hz$.

particolare si noti come, mentre il disallineamento del canale modellato in modo migliore (ossia μ_{\min}) diminuisce sempre in maniera sostanzialmente monotona all'aumentare di L_A , ciò non avviene per μ_{\max} . Più precisamente, per L_A oltre una determinata regione, μ_{\max} cessa di diminuire, assestandosi asintoticamente intorno a un valore μ_{∞} . Si noti anche come la scala usata in figura 1.5 per l'asse delle ascisse sia logaritmica: una rappresentazione lineare, pur espandendo il grafico e rendendone meno intuitiva la lettura, avrebbe accentuato ancora di più il comportamento asintotico. Il comportamento descritto autorizza una doppia considerazione:

- Monitorando un'accuratezza globale nella modellazione del sistema, come ad esempio il disallineamento descritto dalla (1.20), all'aumentare di L_A continuerebbe ad osservarsi un aumento di accuratezza. Ciò ovviamente non è in grado di evidenziare se gli L_A CAP così trovati siano effettivamente CAP “veri”, o soltanto una serie di coefficienti che, nel modello a CAP stimato, modellano accuratamente lo specifico set di N_H risposte impulsive usate come campione (o magari un sottoinsieme di queste ultime);
- L'assestamento di μ_{\max} consente di verificare che, superato un certo valore \bar{L}_A , i *nuovi* poli trovati non introducono un incremento di accuratezza per almeno una delle risposte impulsive usate come campione. Se dunque il campione di RI è scelto in modo adeguato, questo andamento fornisce un'indicazione quantitativa per la scelta dell'ordine efficace \bar{L}_A .

Resta da determinare un criterio operativo per l'individuazione non ambigua dell'ascissa di ginocchio \bar{L}_A nella curva $\mu_{\max}(L_A)$. Per semplicità, in questo

lavoro è stato adottato il criterio euristico

$$\bar{L}_A : \quad \mu_{\max}(\bar{L}_A) = \mu_{\infty} + 1 \text{ dB} \quad (1.22)$$

Le figure 1.7, 1.8, 1.9, 1.10 e 1.11 riassumono i risultati ottenuti ripercorrendo i passi operativi da 1 a 4, elencati alla fine del paragrafo 1.4.2, per le cinque rispettive larghezze di banda scelte. Ogni figura comprende quattro grafici, che riassumono la successione di prove effettuate per determinare sperimentalmente la coppia di ordini efficaci L_A ed L_B . Per ogni valore di f_B :

- Il grafico in alto a sinistra riassume le prove usate per la prima scelta di L_B , come da punto 2 alla fine del par 1.4.2, per un valore di L_A fissato e molto maggiore del valore efficace. E' riportata in rosso la soglia di -40 dB, usata per questa prima scelta di \bar{L}_B , come il minor valore di L_B tale da rendere $\mu_i(\bar{L}_B) < \bar{\mu} \forall i = 1, \dots, N_H$
- Fissato momentaneamente $L_B = \bar{L}_B$, i grafici in alto a destra prima, e in basso a sinistra poi, usano le nuove prove di stima al variare di L_A per determinare \bar{L}_A come descritto al passo 3 del par 1.4.2:
 - Il primo riporta variazioni di L_A su larga scala, permettendo di individuare l'andamento asintotico μ_{∞} per la curva $\mu_{\max}(L_A)$, insieme a un intervallo di ricerca più ristretto per L_A . Sono indicati in rosso i due livelli μ_{∞} e $\mu_{\infty} + 1$ dB, per un uso visivo della (1.22).
 - Il secondo propone la ripetizione delle stesse prove in un intorno dell'ascissa di intersezione tra la curva $\mu_{\max}(L_A)$ e la linea orizzontale $\mu = \mu_{\infty} + 1$ dB, e permette di scegliere \bar{L}_A come il più basso valore di L_A che soddisfi la (1.22).
- L'ultimo grafico, in basso a destra, fissa $L_A = \bar{L}_A$ e ripete le prove del primo grafico per effettuare una nuova (e ultima) scelta di \bar{L}_B , e onorare così il passo 4 del paragrafo 1.4.2.

Una sintesi degli ordini efficaci \bar{L}_A e \bar{L}_B trovati al variare di f_B è presentata in tabella 1.1. La figura 1.6 presenta i risultati graficamente, con della larghezza di banda f_B , nella sua versione normalizzata f_B/f_0 , dove f_0 è a una frequenza *caratteristica* dell'ambiente usato. In questo caso f_0 è stata scelta come la più piccola tra le frequenze modali, ovvero $f_0 = F_{1,1,1}$ data dalla (1.16) per $k_1 = k_2 = k_3 = 1$. Sullo stesso grafico sono riportati i valori teorici per l'ordine CAP, $L_A^T(f_B) = 2 \times n_M(f_B)$, calcolati con entrambe le formule, rispettivamente esatta (1.16) e approssimata (1.17).

Appare evidente che l'ordine efficace trovato è di molto inferiore a quello teorico, come predetto dalle considerazioni del paragrafo 1.2 in merito alle limitazioni dei metodi di stima in generale (e in particolare qui per l'HMK) e da

Tabella 1.1: Larghezze di banda usate e corrispondenti ordini L_A and L_B individuati per, rispettivamente, denominatore comune e singoli numeratori, imponendo un diallineamento di canale massimo di -40 dB. La colonna centrale riporta l'ordine CAP teorico L_A , ottenuto contando le possibili triplette (k_1, k_2, k_3) tali da rendere $F_{k_1, k_2, k_3} < f_B$ (cfr. ...).

f_B (Hz)	f_B/f_0	$2 \times n_M$	\bar{L}_A	\bar{L}_B
250	6.1	68	28	18
333	8.1	144	34	26
500	12.1	434	49	40
667	16.2	962	65	57
1000	24.3	3014	85	89

quelle del paragrafo 1.3.4 rispetto alla densità delle frequenze modali al crescere di f_B . E' inoltre interessante osservare come, in scala doppiamente logaritmica, tale andamento continui ad essere lineare, individuando una dipendenza sperimentale (o forse piuttosto *strumentale*, ovvero legata al metodo di misura) di L_B da f_B secondo una potenza approssimativamente pari ad 1, invece di 3.

Si noti che non è stato presentato alcun paragone con metodi di scelta dell'ordine pre-esistenti, basati unicamente sull'uso dell'accuratezza globale in (1.20). La ragione di questa scelta è che tali approcci permettono di selezionare L_A da un intervallo talmente ampio (a seconda della scelta di L_B) che l'intrinseca ambiguità della scelta stessa non consente un paragone obiettivo con il metodo proposto.

1.5 Conclusioni

Questo primo capitolo ha condotto al primo risultato originale raggiunto da questo lavoro di ricerca: l'individuazione di un nuovo criterio per la selezione di una coppia di ordini (L_A, L_B) nella modellazione basata su CAP. Ciò è avvenuto partendo da un'introduzione a questo tipo di modellazione e da una panoramica sui principali metodi di stima usati in letteratura per il calcolo dei parametri del modello in oggetto: in base al metodo HMK qui adottato questo include gli $N_H \times L_B$ coefficienti degli N_H numeratori di singolo canale, oltre agli L_A coefficienti del denominatore comune. La digressione offerta nel paragrafo 1.3 ha permesso di individuare da una parte i motivi dietro l'uso di modelli basati su CAP, e dall'altra le ragioni dietro le difficoltà di stima dei CAP, per ordini confrontabili a quelli predetti dalle stime teoriche. Delineato infine il motivo per

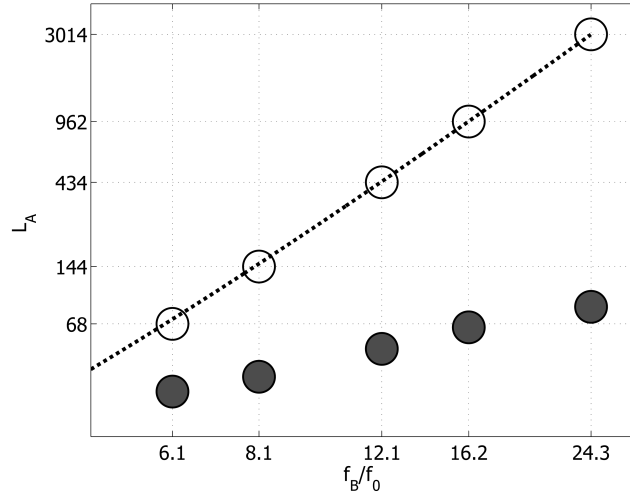


Figura 1.6: \bar{L}_A (Cerchi in grigio pieno) in funzione della larghezza di banda, normalizzata rispetto alle dimensioni ambientali considerate. I cerchi in bianco rappresentano i valori teorici per L_A , come predetto a partire dalla (1.16); la linea tratteggiata si riferisce invece all'espressione approssimata (1.17)

cui i tradizionali metodi di selezione dell'ordine trovano scarsa applicazione in questo caso, è stata introdotta una strategia per selezionare senza ambiguità un ordine efficace, legato insieme all'ambiente da modellare e al metodo di stima dei CAP, e si è osservato in che misura tale ordine differisce tipicamente dalle predizioni teoriche. Tale metodo sarà incontrato nuovamente nel seguito di questo lavoro, e in particolar modo nel capitolo 2, in cui sarà usato per supportare una proposta di nuova architettura per la cancellazione d'eco stereofonica.

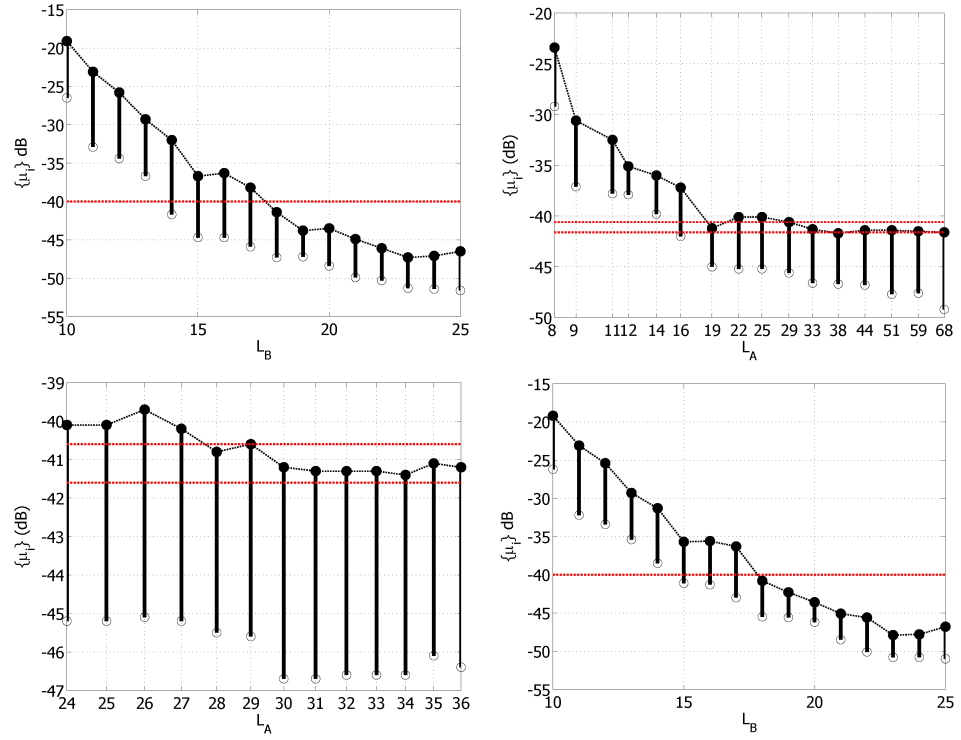


Figura 1.7: $f_B = 250$ Hz (Decimazione 16): Sintesi grafica delle operazioni necessarie alla selezione degli ordini di modellazione. La procedura, da sinistra a destra e dall'alto in basso, è quella descritta nei passi da 1 a 4 alla fine del paragrafo 1.4.2. Il terzo grafico rappresenta simulazioni in un intervallo ristretto del secondo grafico, e porta qui alla selezione di $\bar{L}_A = 28$; entrambi grafici realizzano il terzo passo del metodo. Il primo e il quarto grafico si riferiscono invece alla scelta di L_B , rispettivamente prima (passo 2) e dopo (passo 4) la selezione di \bar{L}_A . Dall'ultima serie di simulazioni deriva in questo caso la scelta $\bar{L}_B = 18$.

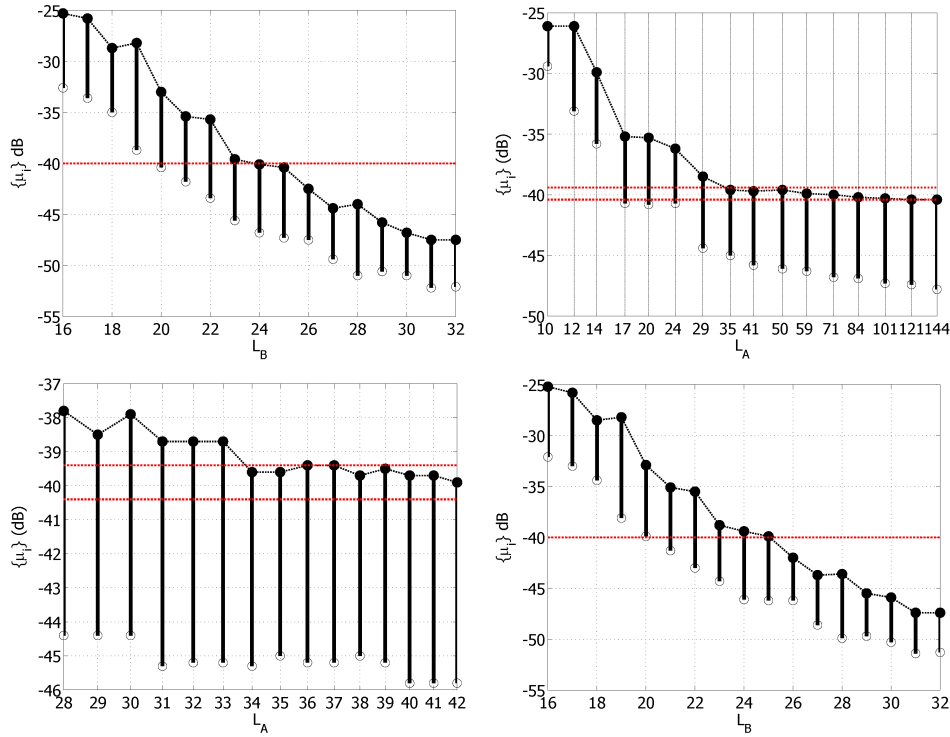


Figura 1.8: $f_B = 333$ Hz (Decimazione 12): Sintesi grafica delle operazioni necessarie alla selezione degli ordini di modellazione. La procedura, da sinistra a destra e dall'alto in basso, è quella descritta nei passi da 1 a 4 alla fine del paragrafo 1.4.2. Il terzo grafico rappresenta simulazioni in un intervallo ristretto del secondo grafico, e porta qui alla selezione di $\bar{L}_A = 34$; entrambi grafici realizzano il terzo passo del metodo. Il primo e il quarto grafico si riferiscono invece alla scelta di L_B , rispettivamente prima (passo 2) e dopo (passo 3) la selezione di \bar{L}_A . Dall'ultima serie di simulazioni deriva in questo caso la scelta $\bar{L}_B = 26$.

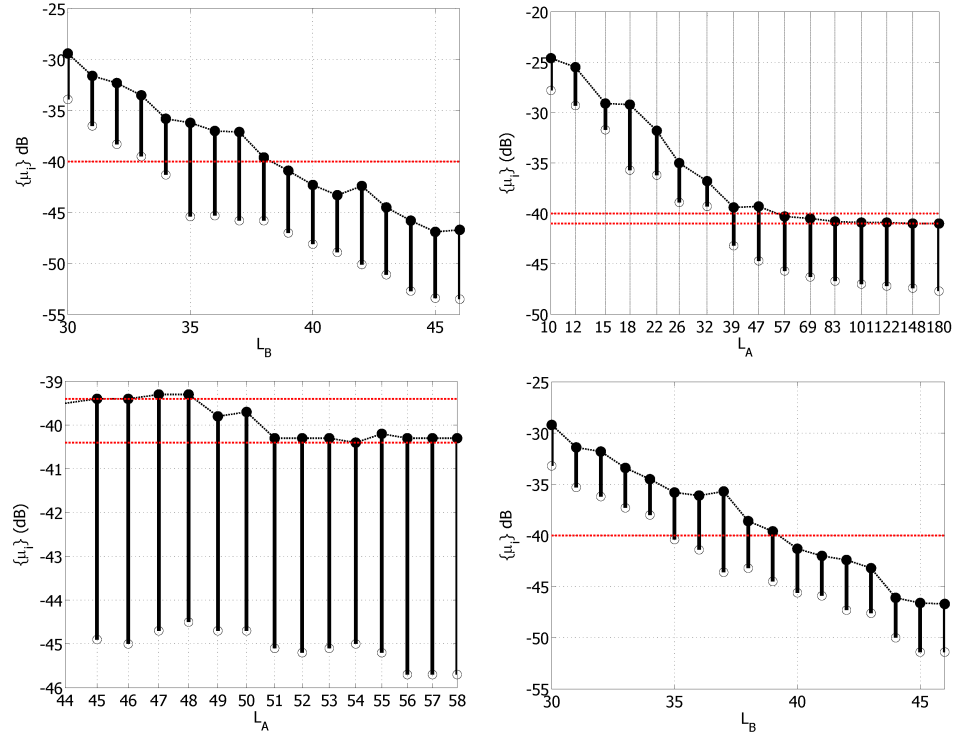


Figura 1.9: $f_B = 500$ Hz (Decimazione 8): Sintesi grafica delle operazioni necessarie alla selezione degli ordini di modellazione. La procedura, da sinistra a destra e dall'alto in basso, è quella descritta nei passi da 1 a 4 alla fine del paragrafo 1.4.2. Il terzo grafico rappresenta simulazioni in un intervallo ristretto del secondo grafico, e porta qui alla selezione di $\bar{L}_A = 49$; entrambi grafici realizzano il terzo passo del metodo. Il primo e il quarto grafico si riferiscono invece alla scelta di L_B , rispettivamente prima (passo 2) e dopo (passo 3) la selezione di \bar{L}_A . Dall'ultima serie di simulazioni deriva in questo caso la scelta $\bar{L}_B = 40$.

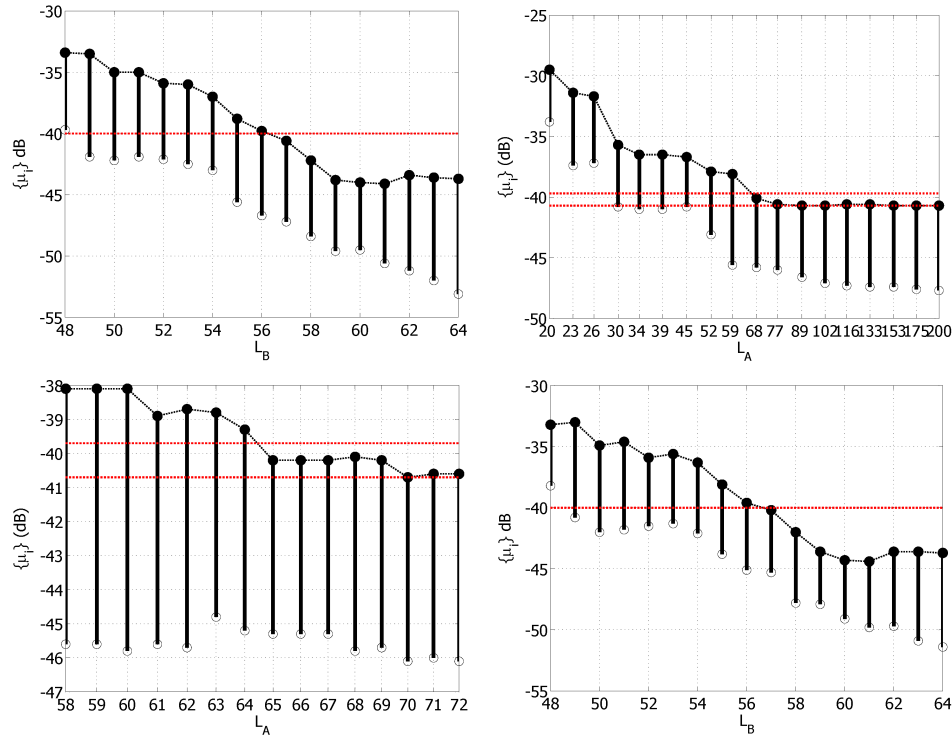


Figura 1.10: $f_B = 667$ Hz (Decimazione 6): Sintesi grafica delle operazioni necessarie alla selezione degli ordini di modellazione. La procedura, da sinistra a destra e dall'alto in basso, è quella descritta nei passi da 1 a 4 alla fine del paragrafo 1.4.2. Il terzo grafico rappresenta simulazioni in un intervallo ristretto del secondo grafico, e porta qui alla selezione di $\bar{L}_A = 65$; entrambi grafici realizzano il terzo passo del metodo. Il primo e il quarto grafico si riferiscono invece alla scelta di L_B , rispettivamente prima (passo 2) e dopo (passo 3) la selezione di \bar{L}_A . Dall'ultima serie di simulazioni deriva in questo caso la scelta $\bar{L}_B = 57$.

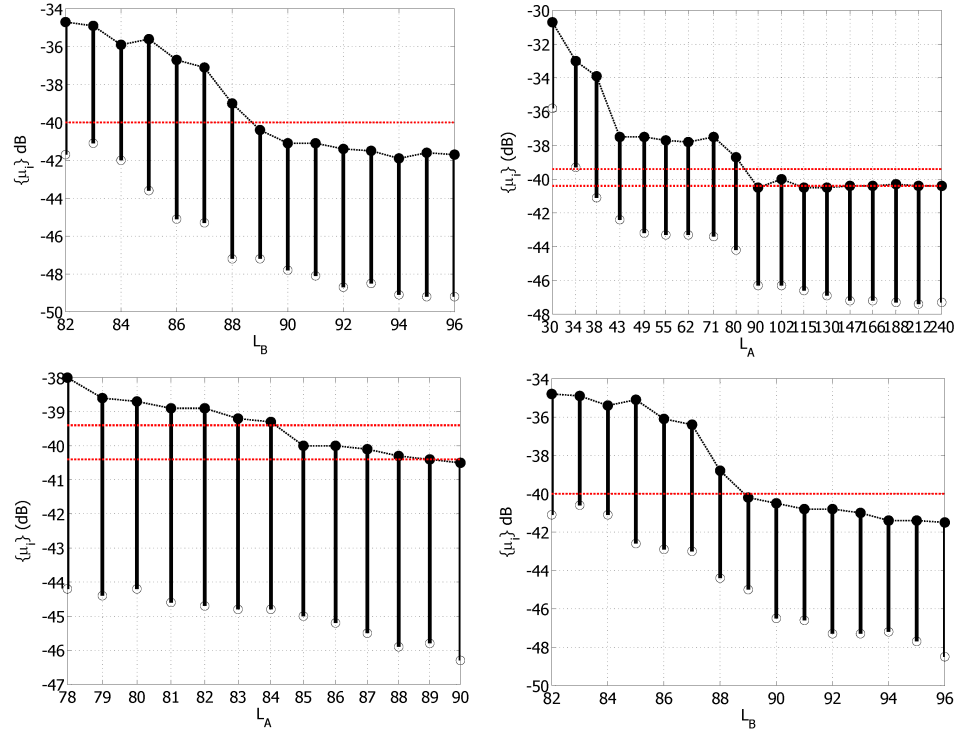


Figura 1.11: $f_B = 1000$ Hz (Decimazione 4): Sintesi grafica delle operazioni necessarie alla selezione degli ordini di modellazione. La procedura, da sinistra a destra e dall'alto in basso, è quella descritta nei passi da 1 a 4 alla fine del paragrafo 1.4.2. Il terzo grafico rappresenta simulazioni in un intervallo ristretto del secondo grafico, e porta qui alla selezione di $\bar{L}_A = 85$; entrambi grafici realizzano il terzo passo del metodo. Il primo e il quarto grafico si riferiscono invece alla scelta di L_B , rispettivamente prima (passo 2) e dopo (passo 3) la selezione di \bar{L}_A . Dall'ultima serie di simulazioni deriva in questo caso la scelta $\bar{L}_B = 89$.

Capitolo 2

La cancellazione d'eco stereofonica con Poli Acustici Comuni

In questo capitolo verrà presentata l'applicazione della modellazione di RTF basata su CAP a un problema che, pur specifico, è allo stesso tempo rappresentativo della vasta classe di problemi adattativi per l'elaborazione numerica in sistemi acustici multicanale. Considerare la cancellazione d'eco stereofonica (o SAEC dall'inglese Stereo Acustical Echo Cancellation) insieme alla modellazione di RTF basata su CAP ha un triplo motivo di interesse:

- Costituire un banco di prova per il nuovo criterio di scelta dell'ordine CAP in ambito adattativo;
- Mostrare un esempio di uso dei CAP in un'applicazione di acustica numerica multicanale;
- Illustrare un esempio in cui l'uso dei CAP offre benefici particolarmente rilevanti, intervenendo nella soluzione di criticità specifiche;

Sarà nel seguito progressivamente sempre più chiaro quale sia il senso associato ad ognuno dei tre punti appena elencati.

In [6], gli inventori dei CAP usano proprio la cancellazione d'eco (in quel caso monocanale) come primo banco di prova utile a dimostrare l'utilità del criterio modellistico introdotto. I risultati cercati in quel contesto sono legati all'esigenza di validazione del modello, in ultima analisi in base alla sua capacità di migliorare le prestazioni di una tipica applicazione adattativa. Nel caso di questo lavoro, quella motivazione è parzialmente superata, in primo luogo perché già coperta in quei primi lavori (sebbene con non chiari criteri per la

scelta dell'ordine), e in secondo luogo perché interessa esplorare applicazioni in cui l'uso dei CAP sia in grado di introdurre vantaggi decisivi. La conoscenza del denominatore comune in contesti generici consente infatti sì di ridurre il numero di gradi di libertà del sistema da stimare (da esempio con un algoritmo adattativo), ma dà comunque per scontata la procedura preventiva di stima a cui si è dedicata grande attenzione nello scorso capitolo, e a cui ancora si concederanno altre riflessioni nel capitolo 4; tale stima è potenzialmente a rischio di consumare risorse in modo da non rendere conveniente, nel bilancio globale, l'adozione del modello a CAP.

Ci sono poi contesti applicativi in cui i vantaggi introdotti sono così importanti da rendere irrilevante il consumo di risorse nella fase di stima: tra queste si pensa qui che meriti di essere collocata la SAEC.

Questa affermazione sarà motivata nel paragrafo 2.2. Perché le osservazioni che si faranno lì siano facilmente comprensibili, nel paragrafo 2.1 è offerta una breve panoramica introduttiva sulla cancellazione d'eco monocanale basata su tecniche adattative. Ciò include una brevissimo richiamo all'algoritmo LMS (Least Mean Squares), con il principale scopo di introdurre un gergo che aiuterà a seguire il filo del seguito dei ragionamenti, ed è seguito da un'introduzione su come la conoscenza del denominatore comune possa essere sfruttata senza perturbare le architetture adattative che usano semplici filtri non ricorsivi.

La SAEC è introdotta e discussa nel paragrafo 2.2, mentre nel paragrafo 2.3 sono presentati i risultati ottenuti in un set-up simulato, del tutto simile a quello usato nel paragrafo 1.4.3.

2.1 Background sulla cancellazione d'eco a singolo canale

2.1.1 Una semplice applicazione per filtri adattativi

La cancellazione d'eco acustica (o AEC, dall'intuitiva versione inglese), costituisce una tipica applicazione di elaborazione numerica adattativa. In un contesto di teleconferenza in viva voce, focalizzando l'attenzione sull'ambiente acustico *ricevente* delineato in figura 2.1, l'eco acustica è quell'effetto indesiderato, per cui al parlante (remoto, in questo caso), viene ritrasmessa una versione filtrata del proprio segnale, a causa di canale acustico aperto tra l'altoparlante ricevente e il microfono locale. Per compensare questo fenomeno, di per sé difficilmente evitabile, è pratica comune introdurre sistemi artificiali per la cancellazione dell'eco acustica. Detto $x[n]$ il segnale ricevuto e riprodotto dall'altoparlante locale, tali sistemi sottraggono dal segnale acquisito dal microfono, $d[n]$, una versione filtrata di $x[n]$. Il sistema filtrante H rappresentato in figura 2.2 ha l'obiettivo di emulare il comportamento della RTF $G(z)$, sotto l'ipotesi

implicita che il canale rappresentato da quest'ultima sia lineare, costante o al più lentamente variabile nel tempo. In particolare, per tenere conto di variazioni del sistema nel tempo (dovute a potenziali cambiamenti di configurazione di altoparlante e microfono nell'ambiente ricevente), il filtro H è reso adattativo, e controllato da un algoritmo che punta alla minimizzazione di una qualche norma del cosiddetto segnale di errore $e[n] = d[n] - y[n]$. $d[n]$ in questo contesto è detto segnale *desiderato*, e $y[n]$ è l'uscita del filtro.

L'architettura del filtro adattativo $H(z)$ è quasi sempre quella cosiddetta *trasversale*, o FIR. La ragione di questa scelta risiede soprattutto nella relativa semplicità degli algoritmi adattativi per filtri FIR, e parzialmente anche nell'immediatezza interpretativa di cui un filtro FIR gode nella modellazione dei sistemi acustici: la risposta impulsiva (pur troncata) del canale modellato. E' utile ricordare che, rispetto alla maggior parte delle applicazioni di stima del canale per sistemi di telecomunicazioni, i filtri FIR in grado di modellare soddisfacentemente sistemi acustici hanno normalmente ordini molto elevati. La ragioni di ciò sono legate da una parte alle frequenze di campionamento usate, e dall'altra ai tempi di propagazione tipici degli ambienti in gioco.

Prima di discutere l'opportunità di introdurre alternative in tal senso basate sul tema centrale di questo lavoro, sembra importante riassumere le conoscenze chiave sui tradizionali algoritmi adattativi per filtri FIR.

La teoria dei filtri adattativi mette a disposizione numerosi algoritmi per controllare la dinamica dei coefficienti del filtro $H(z)$ quando quest'ultimo ha forma FIR [23, 24, 25, 26]. Basti qui ricordare tra gli altri i più semplici e già citati della famiglia LMS, l'oneroso RLS (Recursive Least Squares), o l'APA (Affine Projection Algorithm) [27]. Per ragioni che si ricorderanno tra poco, capita che un algoritmo, frequentemente definito nel dominio del tempo, abbia la sua versione nel dominio della frequenza [28], basata sull'implementazione del filtro FIR con le tecniche standard dette di *overlap and add* o *overlap and save* [29]. Nonostante l'elevato numero di diversi algoritmi disponibili, è abbastanza raro che la teoria riesca a fornire giustificazioni quantitative e dimostrate sull'inci-

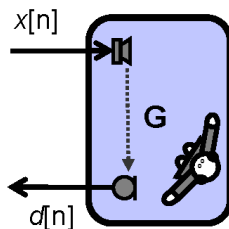


Figura 2.1: Ambiente ricevente, in cui è delineato il contributo detto di eco acustica, che ritrasmette al parlante remoto il proprio stesso segnale, filtrato dal canale ambientale G .

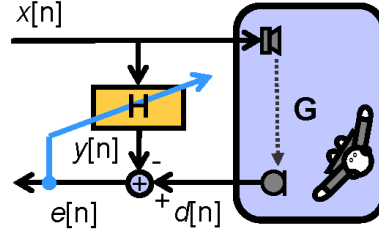


Figura 2.2: Cancellatore artificiale d'eco acustica: Il filtro H è chiamato ad emulare adattativamente il comportamento del canale acustico G , sulla base di un algoritmo che punta a minimizzare una qualche norma del segnale $e[n]$.

denza dei parametri scelti sulle effettive prestazioni dell'algoritmo. Fa eccezione, proprio per la sua semplicità, il classico LMS, e anche in questo caso i risultati delle dimostrazioni sono trovati al prezzo di ipotesi irrealistiche. Nonostante ciò, i risultati trovati per l'LMS sono usati per spiegare qualitativamente, con le dovute precauzioni, anche le prestazioni dinamiche di altri algoritmi. Per questa ragione sembra importante ricordare quali siano questi risultati, e introdurre simboli e terminologia usati nel seguito del lavoro.

2.1.2 Il metodo LMS come riferimento teorico

Come ricordato nel paragrafo precedente, il classico metodo LMS suppone per $H(z)$ una struttura FIR, ovvero

$$H(z) = h_0 + h_1 z^{-1} + \dots + h_{L_H-1} z^{-L_H+1} \quad (2.1)$$

L'uscita $y[n]$ del filtro è calcolata quindi come

$$y[n] = h_n \otimes x[n] = \sum_{k=0}^{L_H-1} h_k x[n-k] = \mathbf{h}^t \cdot \mathbf{x}[n] \quad (2.2)$$

dove l'ultimo membro sfrutta la più compatta notazione vettoriale, insieme alle definizioni:

$$\mathbf{h} = [h_0, h_1, \dots, h_{L_H-1}]^t \quad \text{e} \quad \mathbf{x}[n] = [x[n], x[n-1], \dots, x[n-L_H+1]]^t$$

L'algoritmo adattativo LMS è basato sulla minimizzazione del semplice funzionale quadratico $J(n; \mathbf{h}, \mathbf{x}) = e^2[n]$, e sull'adattamento del filtro \mathbf{h} lungo l'opposto del gradiente $dJ/d\mathbf{h}$. Tralasciando la derivazione analitica, che il lettore può trovare in uno qualsiasi dei testi di riferimento già citati, l'equazione ricorsiva per i coefficienti del filtro in forma vettoriale è la seguente

$$\mathbf{h}[n] = \mathbf{h}[n-1] + \mu e[n] \mathbf{x}[n] \quad (2.3)$$

dove la dipendenza dal tempo associata al vettore dei coefficienti del filtro, $\mathbf{h}[n]$ evidenzia appunto come il valore di questi ultimi cambi, in virtù della precedente equazione, ogni periodo di campionamento. La costante μ , scelta arbitrariamente o ricavata con metodi adattativi aggiuntivi, regola la velocità di aggiornamento dei coefficienti \mathbf{h} , e una sua scelta ottimale è di centrale importanza le prestazioni dell'algoritmo adattativo: valori troppo piccoli di μ conducono ad adattamenti eccessivamente lenti (anche se potenzialmente accurati), mentre valori troppo elevati conducono a scarsa accuratezza del comportamento di $\mathbf{h}[n]$ all'equilibrio, e potenzialmente ad instabilità.

Come accennato nel paragrafo precedente, pur sotto ipotesi irrealistiche è possibile derivare per l'LMS delle dimostrazioni quantitative di grande importanza. Senza entrare nel merito di tali dimostrazioni, se ne cercheranno di ricordare qui le conclusioni più importanti.

Gran parte delle proprietà del filtro adattativo sono legate al problema spettrale della matrice di autocorrelazione degli ingressi, definita come

$$\mathbf{R} = E\{\mathbf{x} \cdot \mathbf{x}^t\} = \begin{pmatrix} r[0] & r[-1] & & r[-L_H + 1] \\ r[1] & r[0] & & \\ & & \ddots & \vdots \\ r[L_H - 1] & & \cdots & r[0] \end{pmatrix}$$

dove $r[i - j] = E\{x[n - i] \cdot x[n - j]\}$.

Si può mostrare [30, 23] che \mathbf{R} ammette diagonalizzazione

$$\mathbf{R} = \mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^t \quad (2.4)$$

essendo \mathbf{Q} la matrice modale unitaria e

$$\mathbf{\Lambda} = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & & \\ \vdots & & \ddots & \vdots \\ 0 & & \cdots & \lambda_{L_H} \end{pmatrix}$$

La distribuzione degli autovalori λ_i è di centrale importanza per la convergenza del filtro LMS. Innanzi tutto per la stabilità del filtro è richiesto che μ soddisfi

$$0 < \mu < \frac{2}{\lambda_{\max}}$$

E' possibile poi mostrare che una versione del filtro ruotata secondo la diagonalizzazione (2.4) converge in modo disaccoppiato, coefficiente per coefficiente, secondo delle costanti di tempo τ_i legate agli autovalori λ_i tramite la relazione $\tau_i \approx \frac{1}{\mu\lambda_i}$. Un'importante conseguenza di ciò è che la convergenza del filtro \mathbf{h} (non ruotato) è legata alla più lenta componente modale, ovvero

$$\tau_h = \frac{1}{\mu\lambda_{\min}}$$

minimizzabile, con la massima scelta ammessa per μ , fino a

$$\min(\tau_h) = \frac{\lambda_{\max}}{2\lambda_{\min}} = \frac{1}{2}\text{cond}(R) \quad (2.5)$$

La (2.5) riassume l'importante conclusione che il tempo caratteristico di convergenza del filtro LMS $\mathbf{h}[n]$ dipende linearmente dal numero di condizionamento $K = \text{cond}(\mathbf{R})$ della matrice di autocorrelazione $L_H \times L_H$ del segnale d'ingresso del filtro, $x[n]$.

Il numero K è evidentemente espressione delle proprietà statistiche di $x[n]$. E' possibile mostrare [25] che, detta $X(f)$ la densità spettrale (o PSD) di $x[n]$, vale

$$K = \frac{\max(S(f))}{\min(S(f))} \quad (2.6)$$

Dunque, segnali a spettro sufficientemente piatto producono una matrice \mathbf{R} ben condizionata, permettendo a \mathbf{h} di convergere rapidamente; al contrario, segnali con grandi escursioni spettrali (per esempio con chiare componenti tonali, o semplicemente filtrati da sistemi selettivi o risonanti) inducono un alto numero di condizionamento per \mathbf{R} , con un conseguente impatto negativo sui tempi di convergenza di $\mathbf{h}[n]$.

Se le proprietà spettrali del segnale d'ingresso nel filtro, $x[n]$ hanno, come appena ricordato, un impatto decisivo sul condizionamento di \mathbf{R} , è importante anche osservare che un ruolo chiave è giocato dalla lunghezza L_H dello stesso filtro LMS. E' infatti possibile mostrare [23] che, a parità di $x[n]$ K è funzione monotona non crescente di L_H , e che quindi, sebbene una maggiore lunghezza del filtro adattativo possa far sperare in una maggiore accuratezza di modellazione, questa scelta ha buone probabilità di incidere negativamente sulle proprietà dinamiche del filtro stesso.

Pur fatte le debite distinzioni, si osserva sperimentalmente che i parametri chiave K e L_H , che controllano l'adattamento del filtro LMS, hanno qualitativamente la stessa incidenza anche per diversi algoritmi adattativi. Gli algoritmi che si dimostrano maggiormente insensibili al problema del malcondizionamento di \mathbf{R} , come il noto RLS, di fatto stimano dinamicamente la matrice \mathbf{R} (o una sua proiezione su un sottospazio a rango ridotto) e la usano per normalizzare selettivamente i diversi modi di convergenza del filtro adattativo: ciò può ridurre la sensibilità rispetto al segnale d'ingresso e alla lunghezza del filtro. D'altra parte l'aumento di robustezza rispetto a K si paga con la complessità dell'algoritmo adattativo, crescente rapidamente con L_H . E' da ricordare anche una spiegazione intuitiva che giustifica l'uso di algoritmi adattativi nel dominio della frequenza: riducendo la banda d'interesse del filtro, e idealmente filtrando separatamente ogni diversa componente in frequenza, la riduzione della larghezza di banda aumenta la probabilità che il rapporto spettrale che definisce K in (2.6) assuma valori contenuti.

In conclusione di questo paragrafo, sembra utile riassumere le difficoltà maggiori incontrate nella cancellazione d'eco monocanale adattativa, sulla scorta della sintesi fatta a proposito dell'algoritmo LMS:

- Il problema centrale è il potenziale elevato valore del numero di condizionamento K della matrice di correlazione dell'ingresso \mathbf{R} ;
- Valori elevati di K possono essere causati da disomogeneità spettrali, come anche da alti valori della lunghezza del filtro adattativo L_H
- Se l'uso del filtro adattativo è inserito nel contesto della cancellazione d'eco, le due principali cause di alti K si riflettono in:
 1. L'uso del parlato come segnale d'ingresso per il filtro adattativo può essere un fattore critico, in particolare per la probabile presenza suoni vocalici;
 2. La lunghezza delle risposte impulsive da stimare con filtri adattativi è un altro grande fattore di criticità, perché come ricordato in precedenza questa è proprio una prerogativa dei sistemi acustici distribuiti.

Nel paragrafo 2.2 si vedrà come la SAEC, pur mantenendo simili criticità di base, sia in realtà caratterizzata principalmente da problemi di altro tipo, di fatto non presenti nel caso dell'AEC monocanale. Il paragrafo successivo introduce invece brevemente all'uso dei CAP nella cancellazione d'eco, a partire dal caso unidimensionale qui introdotto.

2.1.3 L'uso di CAP nell'AEC monocanale

Uno dei motivi generici principali che rende interessante la modellazione basata su CAP risiede nella riduzione del numero dei gradi di libertà (o coefficienti da stimare) nella definizione di un modello di canale acustico. Nel particolare contesto della stima adattativa di canali acustici, di cui l'AEC è un caso particolare, questa attrattiva è ridefinibile in modo più mirato come la diminuzione della lunghezza del filtro adattativo utile a modellare il canale.

Come è noto ed è stato ricordato in precedenza, una delle più rilevanti peculiarità dei canali acustici è la notevole lunghezza delle risposte impulsive che li caratterizzano. Ovviamente ciò si riflette nella lunghezza che un filtro FIR adattativo deve poter onorare, perché nel caso della cancellazione d'eco la risposta impulsiva stimata, rappresentata dal vettore dei coefficienti, modelli in modo sufficientemente accurato il canale obiettivo.

Se, come ricordato nel precedente paragrafo, il fattore di condizionamento K della matrice di correlazione degli ingressi \mathbf{R} è funzione monotona non decrescente della lunghezza L_H del filtro adattativo, è evidente che la possibilità di

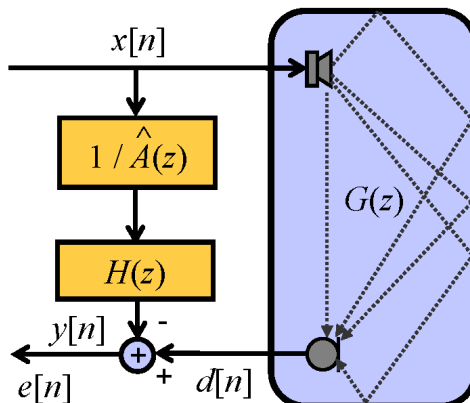


Figura 2.3: Una possibile realizzazione di cancellatore d'eco con il denominatore dei poli comuni: $H(z)$ è il filtro adattativo (per semplicità non è stata usata nessuna simbologia esplicita), e il blocco che lo precede è una funzione di trasferimento costante, contenente unicamente la stima della componente AR comune

disporre del denominatore dei poli comuni, come per esempio indicato in figura 2.3, promette di introdurre notevoli vantaggi. Di fatto una tale architettura non modifica in nulla l'algoritmo che regola l'adattamento del filtro adattativo $H(z)$. Stavolta però, invece di dover stimare l'intero canale $G(z)$, implicitamente assunto della forma $G(z) = B(z)/A(z)$, $H(z)$ è unicamente chiamato a fornire, all'equilibrio, una stima $\hat{B}(z)$ del numeratore $B(z)$. In termini di ciò che conta per le prestazioni del filtro adattativo, a parità di accuratezza di modellazione globale la nuova lunghezza di \mathbf{h} sarà potenzialmente assai inferiore a quella originale, con grandi possibilità di riduzione del numero di condizionamento K della matrice di correlazione.

A dispetto dell'apparente semplicità, la soluzione appena descritta nasconde un non evidente tranello. Si ricorderà che il numero di condizionamento K , oltre ad essere dipendente dalla lunghezza del filtro adattativo, è innanzi tutto interpretabile come il rapporto tra i valori estremi dello spettro del segnale d'ingresso, così come espresso dalla (2.6). Nell'AEC modificato di figura 2.3, il segnale d'ingresso del filtro adattativo non è più $x[n]$, ma la sua versione filtrata dal sistema costante $1/\hat{A}(z)$, qui detta $z[n]$, essendo $\hat{A}(z)$ la stima del denominatore dei poli comuni. Per ragioni che dovrebbero essere già parzialmente chiare dal paragrafo 1.3, il sistema $1/\hat{A}(z)$ ha di per sé una risposta in frequenza contraddistinta da numerosissimi picchi di risonanza, tanto più pronunciati quanto minori sono le perdite per assorbimento che caratterizzano il sistema acustico scelto. In conseguenza di ciò, la nuova matrice di autocorrelazione

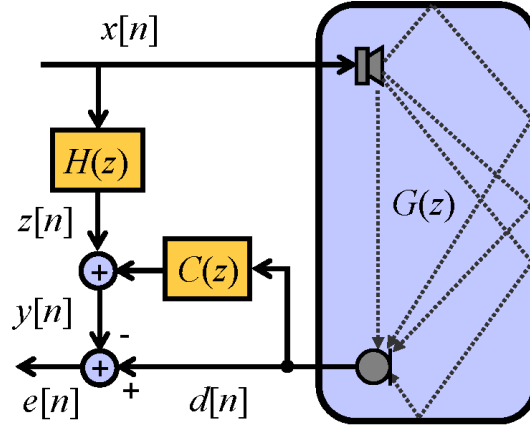


Figura 2.4: Una più conveniente realizzazione del pre-filtraggio con il denominatore dei poli comuni, alternativa a quella descritta in figura 2.3. Qui è il blocco $C(z) = 1 - \hat{A}(z)$ ad introdurre nella cancellazione d'eco la conoscenza dei CAP. Si noti come in questo caso anche quest'ultimo filtro fisso ha un'implementazione non ricorsiva.

dell'ingresso

$$R_{zz} = E\{\mathbf{z} \cdot \mathbf{z}^t\} \quad \text{con} \quad \mathbf{z} = [z[n], \dots, z[n - L_H + 1]]^t$$

ha buone probabilità di risultare malcondizionata, per quanto piccola possa essere la nuova lunghezza L_H del filtro adattativo ridotto. Come è ovvio, questo effetto compenserebbe negativamente i vantaggi introdotti dalla conoscenza di $\hat{A}(z)$, e per evitarlo l'idea a monte dell'architettura di figura 2.3 è praticamente realizzata con la struttura descritta in figura 2.4. Per capire in che modo le due strutture possano essere ritenute funzionalmente equivalenti, si consideri in quest'ultimo caso l'espressione del segnale d'errore $e[n]$, nel dominio z e in funzione del segnale d'ingresso X .

$$E(z) = [1 - C(z)] \cdot G(z) \cdot X(z) - H(z) \cdot X(z) \neq 0$$

Se l'algoritmo adattativo per $H(z)$ punta alla minimizzazione dell'energia di $e[n]$, in condizioni di adattamento raggiunto varrà

$$E(z) \simeq 0 \quad \forall X(z)$$

e quindi, omettendo per semplicità la dipendenza in z

$$(1 - C) \cdot G \simeq H$$

Sostituendo a G la sua forma ARMA in funzione del denominatore dei CAP, B/A , e ricordando che si desidera che H produca la stima del numeratore,

definita come \hat{B} , si ottiene

$$(1 - C) \cdot \frac{B}{A} \simeq \hat{B}$$

Ammettendo che la stima \hat{B} debba essere quanto più possibile fedele al suo obiettivo B , e ricordando che il filtro C è definito artificialmente, si arriva alla semplice espressione

$$C(z) = 1 - \hat{A}(z) \quad (2.7)$$

dove $\hat{A}(z)$ è ancora una volta la stima del denominatore dei CAP, $A(z)$.

Si noti che la struttura così definita offre notevoli vantaggi rispetto alla sua precedente formulazione:

- La matrice di correlazione \mathbf{R} che regola le prestazioni dell'algoritmo adattativo è ancora costruita con il segnale d'ingresso originario $x[n]$, e non con la sua versione filtrata $z[n]$; le dimensioni della matrice di correlazione saranno tuttavia inferiori, così come desiderato con l'introduzione della conoscenza del denominatore dei CAP.
- La conoscenza del denominatore comune è introdotta con l'uso del filtro costante $C(z)$ che ha una semplice struttura FIR, con i coefficienti derivabili dalla (2.7):

$$\mathbf{c} = [c_1, c_2, \dots, c_{L_A}] = [-a_1, -a_2, \dots, -a_{L_A}] \quad (2.8)$$

Qui L_A è ovviamente l'ordine della stima del denominatore comune.

L'uso dei CAP nella formulazione appena descritta fu già introdotta in [6], senza che però gli autori ne fornissero una spiegazione diretta. I risultati presentati in questo paragrafo non devono quindi considerarsi originali, sebbene una spiegazione dettagliata, a conoscenza dell'autore, non fosse stata esplicitamente presentata.

Come già accennato, l'uso dei CAP per l'AEC monocanale fu usato in [6] proprio come banco di prova dell'efficacia e della validità della stessa modellazione basata su CAP. I risultati lì presentati hanno mostrato un confronto tra l'architettura di figura 2.4 e un AEC canonico con il solo filtro FIR (come ad esempio in figura 2.2), a parità di algoritmo adattativo scelto per il filtro H e a parità di energia dell'errore all'equilibrio, o piuttosto della sua più comune versione normalizzata detta ERLE (Acronimo dell'inglese *Echo Return Loss Enhancement*). In particolare, le prove lì fatte hanno seguito i seguenti passi:

1. Scelta di un ambiente simulato, e selezione di un certo numero N_H di risposte di canale campione.
2. Stima del denominatore comune $\hat{A}(z)$, con a monte selezione degli ordini L_A ed L_B (seppur con criteri di scelta dell'ordine non del tutto chiari, come ricordato nel paragrafo 1.4).

3. Realizzazione di un cancellatore d'eco con CAP, secondo lo schema di figura 2.4 e con un filtro adattativo di lunghezza L_B . In questo caso usando un nuovo canale rispetto a quelli già selezionati come campioni, e precedentemente usati per la stima di $A(z)$.
4. Realizzazione di un AEC tradizionale per lo stesso canale, con una lunghezza L_H per il filtro FIR adattativo tale da rendere l'ERLE all'equilibrio pari a quello del punto 3.
5. Confronto tra L_B ed L_H , e tra i tempi di convergenza caratteristici nei due casi (indici diretti dei guadagni di prestazione ottenuti con la riduzione della dimensionalità di \mathbf{R} , e conseguentemente del suo numero di condizionamento K).

Haneda e gli altri autori furono in grado di osservare in [6] un'importante riduzione sia del numero di coefficienti adattativi necessari che del tempo caratteristico di convergenza, usando il *Normalized LMS* (o NLMS) come algoritmo adattativo in entrambi i casi. Le ragioni di tali miglioramenti, sostanzialmente attese in base a quanto argomentato fin qui, costituivano lì una prova sperimentale dell'efficacia della modellazione basata su CAP. La scelta dell'NLMS era lì ottimale per mettere in luce l'efficacia dell'introduzione dei CAP, a causa dell'estrema sensibilità della famiglia di algoritmi LMS al malcondizionamento della matrice di correlazione (come ricordato nel paragrafo 2.1.2). Sembra anche importante notare che i risultati lì mostrati non erano peraltro quantitativamente di enorme rilievo, probabilmente proprio a causa della mancata definizione di un criterio non ambiguo per la scelta degli ordini L_A ed L_B .

Anche se nel presente lavoro è stato introdotto un nuovo criterio di scelta dell'ordine, in grado di produrre risultati più decisivi di quelli a suo tempo trovati in [6], si è deciso qui di non dare troppo spazio a questa semplice applicazione, e di dedicare maggiore attenzione al più interessante caso della cancellazione d'eco stereofonica che sarà introdotta nel prossimo paragrafo. L'approccio metodologico usato in [6], è stato comunque in larga parte mantenuto per le prove descritte nel paragrafo 2.3.

2.2 La cancellazione d'eco stereofonica

2.2.1 Descrizione del problema

Il contesto applicativo in cui si definisce la cancellazione d'eco stereofonica (d'ora in poi SAEC) è assai simile a quello dell'AEC. Questa volta però il sistema telefonico in viva-voce è supposto a doppio canale: si assume che il contenuto sia prelevato presso il parlante remoto (nell'ambiente di *trasmissione*) con

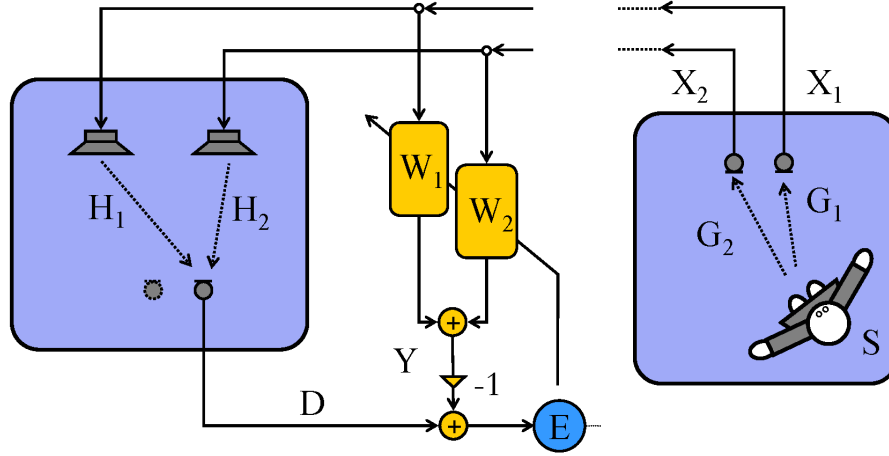


Figura 2.5: Esempio di sistema telefonico in viva-voce stereofonico, con il relativo sistema di cancellazione dell'eco.

una coppia di microfoni invece che con uno solo. Specularmente, il segnale stereofonico è riprodotto a destinazione (nell'ambiente di *ricezione*) mediante una coppia di altoparlanti, e ugualmente nell'altro senso. Le ragioni per una simile complicazione sono molteplici: basti pensare al miglioramento della descrizione spaziale del parlante implicitamente codificata nel segnale sonoro. Uno schema del sistema appena descritto è presentato in figura 2.5.

A causa dell'aumento dei simboli necessari alla descrizione del problema, i filtri sono stati qui associati nel dominio z al simbolo W_i ($i = 1, 2$). Ognuno dei due W_i è chiamato a stimare una singola RTF H_i , definita dall'accoppiamento tra l' i -esimo altoparlante nell'ambiente di ricezione e uno solo dei due microfoni: il sistema descritto è usato per rimuovere l'eco acustica da uno dei due segnali acquisiti nell'ambiente di ricezione, e andrà idealmente replicato integralmente per il secondo microfono.

Sempre nel dominio z , si chiami X_i l' i -esimo segnale ricevuto. Per ragioni che verranno chiarite tra breve, nel caso della SAEC riveste una centrale importanza anche la definizione di come gli X_i sono generati nell'ambiente di trasmissione, dove l'accoppiamento tra la sorgente S e i due microfoni è descritto mediante le due rispettive RTF G_1 e G_2 .

Per una migliore definizione del problema adattativo, si introducono i due seguenti vettori di dimensione $(L_W \times 1)$, essendo L_W la lunghezza di ciascuno dei filtri FIR W_i :

$$\mathbf{x}_i[n] = [x_i[n], x_i[n-1], \dots, x_i[n-L_W+1]]^t \quad i = 1, 2$$

Per quanto detto vale ovviamente

$$x_i[n] = g_i[n] \otimes s[n] \quad i = 1, 2 \quad (2.9)$$

Se, similmente a quanto visto nel paragrafo, 2.1.2, si rappresentano anche le risposte impulsive in notazione vettoriale, si ottiene

$$\begin{aligned} \mathbf{g}_i &= [g_i[n], g_i[n-1], \dots, g_i[n-L_G+1]]^t \\ \mathbf{h}_i &= [h_i[n], h_i[n-1], \dots, h_i[n-L_H+1]]^t \quad \text{per } i = 1, 2 \\ \mathbf{w}_i &= [w_i[n], w_i[n-1], \dots, w_i[n-L_W+1]]^t \end{aligned} \quad (2.10)$$

La conclusione che la soluzione cercata tende a un valore ottimo congiunto per la coppia di filtri adattativi, attraverso la minimizzazione dell'energia dell'unico segnale di errore disponibile

$$e[n] = d[n] - y[n] = d[n] - \mathbf{w}^t \cdot \mathbf{x} \quad (2.11)$$

essendo $d[n]$ il segnale acquisito dal microfono considerato e $y[n]$ la somma delle uscite dei due filtri adattativi. L'ultimo membro della (2.11) va letto alla luce delle definizioni

$$\begin{aligned} \mathbf{w} &= [\mathbf{w}_1^t, \mathbf{w}_2^t]^t \\ \mathbf{x} &= [\mathbf{x}_1^t, \mathbf{x}_2^t]^t \end{aligned} \quad (2.12)$$

Nonostante la forma di $y[n]$ in (2.11) assomigli in tutto e per tutto alla (2.2), il comportamento della SAEC differisce notevolmente dai risultati di singolo canale riassunti nel paragrafo 2.1. In particolare, come mostrato in [31] tra i primi contributi in letteratura, le prestazioni adattative del sistema in questo caso dipendono da una nuova matrice di correlazione

$$\mathbf{R} = E\{\mathbf{x} \cdot \mathbf{x}^t\} = E\left\{\begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{bmatrix} \cdot \begin{bmatrix} \mathbf{x}_1^t & \mathbf{x}_2^t \end{bmatrix}\right\} = \begin{pmatrix} \mathbf{R}_{11} & \mathbf{R}_{12} \\ \mathbf{R}_{21} & \mathbf{R}_{22} \end{pmatrix} \quad (2.13)$$

dove $\mathbf{R}_{ij} = E\{\mathbf{x}_i \cdot \mathbf{x}_j^t\}$. La matrice \mathbf{R} appena definita è spesso detta *matrice di autocorrelazione stereofonica*. Come verrà riassunto nel prossimo paragrafo, le proprietà spettrali di \mathbf{R} sono dipendenti dall'intera architettura di figura 2.5, e condizionano le prestazioni della SAEC adattativa secondo modalità nuove rispetto all'AEC di singolo canale.

2.2.2 Background analitico

Problematicità specifiche

La formulazione stessa del problema adattativo stereofonico rende il funzionamento e le relative prestazioni della SAEC notevolmente diversi rispetto all'AEC monocanale. Come può intuirsi dalle osservazioni fatte nel paragrafo 2.1.2, la chiave del problema è sempre il condizionamento della matrice di correlazione caratterizzante il sistema: in questo caso la matrice di correlazione stereofonica definita in (2.13). Se fin qui il parallelismo con il caso monocanale è mantenuto, ciò in cui la SAEC richiede diverse riflessioni sono le cause che regolano il condizionamento di \mathbf{R} . Si ricorderà che, per l'AEC monocanale, i potenziali fattori che influenzano negativamente il condizionamento di \mathbf{R} sono rintracciabili

da una parte nella coloritura spettrale del segnale riprodotto dall'altoparlante nell'ambiente di ricezione, e dall'altra dalla lunghezza del filtro adattativo. Quest'ultima, a sua volta, ha un'evidente dipendenza dai tempi di riverbero caratteristici del canale da stimare: determinata una lunghezza efficace della risposta impulsiva di canale (in buona sostanza proporzionale al tempo di riverbero della risposta stessa), al filtro adattativo è richiesto di avere una lunghezza non molto minore di quest'ultima. In caso contrario la stima del canale può difficilmente essere accurata, e quindi a sua volta l'energia del segnale di errore difficilmente conducibile a valori sufficientemente piccoli.

Nel caso della SAEC, \mathbf{R} in (2.13) è costituita da due termini diagonali R_{11} e R_{22} , rappresentanti semplici matrici di correlazione relative a ciascuno dei due rispettivi segnali ricevuti, e da due termini *incrociati*, R_{12} e R_{21} , rappresentanti matrici di cross-correlazione tra i due diversi segnali $x_1[n]$ e $x_2[n]$. Questi due termini extra-diagonali contribuiscono negativamente al condizionamento di \mathbf{R} , tanto più quanto maggiore è la dipendenza statistica tra i due vettori \mathbf{x}_1 e \mathbf{x}_2 . Uno dei primi moderni contributi alla comprensione del problema è offerto da Benesty, Morgan e Sondhi in un lavoro del 1998 [31]: gli autori forniscono lì una prima dettagliata dimostrazione delle cause del malcondizionamento di \mathbf{R} nella SAEC, prendendo i passi da osservazioni presentate in precedenti lavori [32, 33]. Sebbene non si ritenga essenziale presentare qui i dettagli delle dimostrazioni offerte, sembra comunque rilevante cercare di delineare i punti chiave dei ragionamenti fatti in [31] e altrove: una comprensione intuitiva delle cause dei principali problemi legati alla SAEC permetterà a sua volta una migliore comprensione delle proposte presentate nel paragrafo 2.2.3 rispetto all'uso dei CAP in questo contesto.

Il metodo seguito in [31] parte dall'assumere che tutte le risposte impulsive dei canali coinvolti possano ragionevolmente esprimersi attraverso semplici filtri numerici FIR di lunghezza limitata secondo le (2.10). Sebbene tale rappresentazione possa apparire limitante (vista l'illimitatezza temporale intrinsecamente associata alla risposta impulsiva di ogni sistema fisico), le descrizioni che ne risultano rendono l'interpretazione del problema di più semplice comprensione: le lunghezze dei modelli FIR rappresentano le lunghezze efficaci dei sistemi fisici corrispondenti, e una tale rappresentazione schematica rende più semplice una descrizione del problema sulla base di tali parametri.

I ragionamenti lì fatti danno per scontato una formulazione adattativa basata sul classico algoritmo *Recursive Least Squares* (RLS), qui non ricordato esplicitamente, ma rintracciabile in qualunque testo specialistico [23, 24, 25, 26]. La principale ragione di questa scelta è che l'RLS introduce una stima dinamica della matrice di correlazione degli ingressi in modo esplicito nelle equazioni

dell'algoritmo, e pertanto alcune conclusioni sono accessibili più rapidamente seguendo tale approccio. Così come ricordato nel presente lavoro prima dei richiami sull'LMS, le conclusioni raggiunte sono da ritenersi comunque valide in modo generico, e pertanto qualitativamente estendibili a qualunque algoritmo adattativo.

Il primo importante punto chiave dei ragionamenti presenti in letteratura sulla SAEC consiste nel quantificare la dipendenza statistica tra i due segnali x_1 e x_2 , per capire come ciò incida nella forma di \mathbf{R}_{ij} per $i \neq j$ e in ultima analisi di \mathbf{R} in (2.13). x_1 e x_2 hanno una dipendenza statistica intrinseca dovuta al fatto di essere entrambi effetto dello stesso segnale $s[n]$ in base alla (2.9). I due segnali mantengono quindi una reciproca dipendenza lineare, espressa nel dominio della frequenza da [31].

$$G_1(f)X_2(f) = G_2(f)X_1(f) \quad (2.14)$$

Detta L_G la lunghezza di ciascuna delle due risposte impulsive dei canali di trasmissione \mathbf{g}_1 e \mathbf{g}_2 , per semplicità assunte dello stesso ordine, ed L_W la lunghezza di ognuno dei due filtri adattativi, vale la seguente

Osservazione 1 [31]. La matrice di correlazione stereofonica \mathbf{R} è singolare se $L_W > L_G$.

La singolarità di \mathbf{R} non solo determina l'impossibilità della convergenza di un approccio adattativo, ma addirittura nega l'esistenza di un'unica soluzione per il problema di Wiener. Questa circostanza sembra a priori facilmente aggirabile, selezionando una lunghezza L_W per i filtri adattativi sufficientemente ridotta. Ciononostante una considerazione collaterale, sempre derivata in [31], incoraggia ad evidenziare anche il seguente risultato:

Osservazione 1a. Anche per $L_W < L_G$, il numero condizionamento di \mathbf{R} cresce al diminuire dell'energia delle code formate degli ultimi $L_G - L_W$ campioni di \mathbf{g}_1 e \mathbf{g}_2 .

Riassumendo qualitativamente le prime conclusioni appena ricordate, è possibile affermare che la dipendenza statistica tra i due segnali d'ingresso del sistema adattativo (intrinsecamente presente per l'origine stessa dei due canali) mina il buon condizionamento di \mathbf{R} , a meno che L_W (lunghezza dei filtri adattativi) non sia molto minore di L_G (lunghezza efficace delle risposte impulsive di trasmissione).

Ciò che si è presentato genericamente come dipendenza statistica è riconducibile quantitativamente alla cosiddetta funzione di coerenza [31, 34] tra i due

segnali d'ingresso, definita da:

$$\gamma(f) = \frac{S_{x_1 x_2}(f)}{\sqrt{S_{x_1 x_1}(f) \cdot S_{x_2 x_2}(f)}} \quad (2.15)$$

dove

$$\begin{aligned} S_{x_i x_k}(f) &= \sum_{\tau=-\infty}^{+\infty} E \{x_i[n]x_k[n-\tau]\} e^{-j2\pi f\tau} \\ &= \sum_{\tau=-\infty}^{+\infty} R_{x_i x_k}(\tau) e^{-j2\pi f\tau} \quad \text{per } i, j = 1, 2 \end{aligned} \quad (2.16)$$

rappresenta la PSD (incrociata) per la coppia i, j . La funzione di coerenza $S_{x_i x_k}(f)$ definita in (2.15) fornisce appunto una misura della dipendenza statistica tra i due ingressi. In particolare, $|\gamma| = 1$ indica una completa dipendenza, in mancanza della quale vale la generica disuguaglianza $0 < |\gamma| < 1$.

Un primo importante risultato [31] è che gli autovalori di \mathbf{R} sono limitati dal basso da

$$1 - |\gamma(f)|^2$$

ossia l'esistenza di un valore f per cui $|\gamma(f)| = 1$ renderebbe singolare \mathbf{R} . Espressioni analitiche, qui non riportate per brevità, sono state derivate anche per la dipendenza diretta del numero di condizionamento di \mathbf{R} da $\gamma(f)$ [34].

Un altro importante problema legato alla lunghezza L_W dei filtri adattativi \mathbf{w}_1 e \mathbf{w}_2 riguarda la relazione con la lunghezza efficace L_H delle risposte impulsive dei due canali acustici nell'ambiente di ricezione, ovvero \mathbf{h}_1 e \mathbf{h}_2 .

Intuitivamente, è facilmente comprensibile che L_W debba essere comparabile con L_H perché una buona stima dei due rispettivi canali sia possibile. In caso contrario \mathbf{w} potrà difficilmente tendere ad \mathbf{h} , con conseguenti scarsi risultati nella minimizzazione dell'energia dell'errore $e[n]$. Usando un gergo più appropriato alla cancellazione d'eco, ciò significherebbe un alto valore per il *Mean Squared Error* (MSE) o simmetricamente un basso valore dell'ERLE.

Osservazione 2. Per $L_W < L_H$ i filtri adattativi non convergono verso le risposte impulsive corrispondenti.

In realtà tale considerazione, facilmente intuibile e valida anche nel caso unidimensionale, nasconde nella SAEC implicazioni ben più problematiche. Infatti se anche L_W non è molto minore di L_H , sono ugualmente raggiungibili valori di MSE (o ERLE) apparentemente soddisfacenti. I vettori a cui tendono \mathbf{w}_1 e \mathbf{w}_2 , però, in questo caso differiscono sostanzialmente dai vettori obiettivo \mathbf{h}_1 e \mathbf{h}_2 . In altri termini la soluzione del problema di Wiener non è unica, e anzi dipende dai due canali \mathbf{g}_1 e \mathbf{g}_2 . Questo significa che se il parlante remoto

cambia posizione nell'ambiente di trasmissione, la cancellazione d'eco conseguita con l'adattamento cessa di essere valida.

Tale problema è ben quantificato con la definizione di un altro importante parametro, utile, come l'MSE o l'ERLE, a descrivere le prestazioni della cancellazione d'eco, e particolarmente utile nel caso della SAEC:

$$\mu[n] = \frac{\|\mathbf{w}[n] - \mathbf{h}\|_2^2}{\|\mathbf{h}\|_2^2} \quad (2.17)$$

$\mu[n]$, definito come il *disallineamento* della coppia di filtri adattativi, rappresenta la norma quadratica, normalizzata rispetto al riferimento, tra i vettori dei filtri e le risposte impulsive dei rispettivi canali, al variare del tempo. Il disallineamento consente di valutare non semplicemente la cancellazione d'eco in sé (ovvero la minimizzazione del segnale di errore) ma soprattutto la convergenza dei filtri verso una soluzione unica, e indipendente dalla posizione del parlante nell'ambiente di trasmissione. E' opportuno osservare che il calcolo di $\mu[n]$ è estremamente utile nelle derivazioni teoriche, oltre essere facilmente accessibile in verifiche sperimentali che sfruttino ambienti simulati; in caso contrario il vettore \mathbf{h} è difficilmente conoscibile, e di conseguenza il valore di $\mu[n]$ non è calcolabile. E' stato ampiamente mostrato come, in situazioni realistiche per cui $L_W < L_H$, il disallineamento $\mu[n]$ sia diretta funzione del condizionamento di \mathbf{R} , e in ultima analisi della funzione di coerenza definita in (2.15).

Sembra opportuno riassumere le ultime considerazioni fatte nella seguente osservazione:

Osservazione 2a. Se non è soddisfatta $L_W \simeq L_H$, la minimizzazione del segnale di errore non fa convergere i filtri \mathbf{w} verso \mathbf{h} , ma verso soluzioni dipendenti da \mathbf{g}_1 e \mathbf{g}_2 . Tale effetto è tanto più pronunciato quanto più è alto il numero di condizionamento di \mathbf{R} .

Ripercorrendo i risultati riassunti fin qui, dovrebbe apparire chiaro che le richieste per un buon funzionamento della SAEC impongono condizioni palesemente contraddittorie sulla lunghezza L_W dei filtri adattativi. In generale, infatti, i due ambienti di trasmissione e di riflessione devono considerarsi acusticamente equivalenti per simmetria, e quindi è ragionevole, se no doveroso, assumere $L_G \simeq L_H = L$. In tali ipotesi le osservazioni 1 e 2, pur nella loro schematicità, fanno emergere in pieno la contraddizione, richiedendo in buona sostanza che la scelta di L_W sia tale da soddisfare

$$L_W : (L_W \ll L) \wedge (L_W \simeq L)$$

Le osservazioni 1a e 2a, pur in modo meno schematico, non fanno altro che ribadire la stessa coppia di condizioni conflgenti. La conclusione è che, così come è stato appena presentato, il problema della cancellazione d'eco stereofonica

è intrinsecamente malposto.

Soluzioni esistenti

La causa dei problemi della SAEC è dunque il malcondizionamento intrinseco della matrice di correlazione stereofonica \mathbf{R} , presente se non vale $L_W \ll L_G$, e tanto più pronunciato quanto più $|\gamma(f)|$ è vicino all'unità.

Il modo in cui diversi autori hanno in passato cercato di aggirare questa criticità è stato sempre centrato sul tentativo di pre-processare la coppia stereofonica dei segnali d'ingresso x_1 e x_2 . Obiettivi (e vincoli) di tali strategie possono riassumersi nei due seguenti punti

1. Disaccoppiare statisticamente i segnali d'ingresso in pre-ricezione, usando due nuove versioni \tilde{x}_1 e \tilde{x}_2 sia come segnali riprodotti che come ingressi dei filtri adattativi;
2. Assicurare che le trasformazioni $\tilde{x}_i = \mathcal{T}[x_i]$, per $i = 1, 2$ non alterino significativamente i due segnali d'ingresso dal punto di vista percettivo, in particolar modo per ciò che riguarda la caratterizzazione spaziale intrinsecamente codificata nella coppia di segnali, che costituisce una delle principali ragioni per l'uso di un sistema di teleconferenza stereofonica.

Nei limiti di cautela imposti dal secondo punto, le trasformazioni proposte in letteratura puntano ovviamente ad allontanare il più possibile dal valore unitario il modulo della funzione di coerenza, $|\gamma(f)|$, per la totalità dello spettro usato. Così facendo è possibile recuperare un buon condizionamento di \mathbf{R} anche per valori ragionevolmente elevati di L_W , sufficienti a loro volta a scongiurare la convergenza verso soluzioni fittizie, ovvero a produrre una buona minimizzazione del disallineamento $\mu[n]$, in tempi relativamente contenuti.

La prima classe di soluzioni importanti da citare è quella che introduce procedure lineari per il disaccoppiamento dei due segnali. Shimauchi a Makino da una parte [35], e Sondhi et al. dall'altra [32] propongono sistemi di filtri adattativi per generare due nuovi segnali d'ingresso, imponendo per quest'ultimi il vincolo di correlazione minima o una qualche forma di ortogonalità geometrica [36]; i risultati riportati variano a seconda dei casi, ma non sembrano promettere un miglioramento importante nella convergenza di $\mu[n]$. Una soluzione diversa, dall'interpretazione intuitiva nel dominio della frequenza, consiste nel filtrare i due rispettivi segnali con diversi filtri a pettine, tra loro interlacciati, in modo che ognuno dei due segnali pre-filtrati abbia dei massimi dello spettro per quei valori di f in cui l'altro ha i punti di minimo [32, 33]. Tale tecnica, che prende in prestito un noto stratagemma usato nella post-produzione audio per generare segnali stereofonici fittizi da tracce monofoniche soffre di due fondamentali

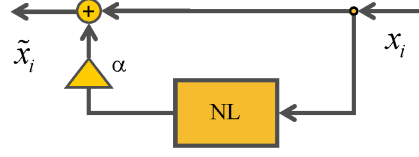


Figura 2.6: Schema di riferimento per il pre-processamento non lineare dei segnali d'ingresso dei filtri adattativi nella SAEC: NL rappresenta il generico blocco non lineare, ed α il parametro di dosaggio. L'orientamento delle frecce è in accordo con lo schema di figura 2.5.

problemi: in primo luogo il disaccoppiamento introdotto non allontana sufficientemente dall'unità la quantità $|\gamma(f)|$; in aggiunta a ciò, l'alterazione percettiva introdotta, apparentemente tollerabile sopra i 1000 Hz, produce artefatti che compromettono una corretta discriminazione spaziale al di sotto di tale soglia. Per ovviare a quest'ultimo inconveniente, è stata anche presentata un'architettura che usa i filtri a pettine per $f > 1$ kHz, e ricorre ai metodi lineari citati tra breve per frequenze al di sotto di tale soglia [33]; un'ulteriore soluzione ibrida è quella presentata in [37].

Un'altra famiglia di metodi lineari usati per disaccoppiare x_1 e x_2 consiste nel sommare ad entrambi due diversi segnali casuali reciprocamente scorrelati. In tale contesto rivestono grande importanza metodi che usano criteri percettivi, per sagomare lo spettro del rumore additivo in modo da minimizzarne l'udibilità [38, 39]. Apparentemente queste ultime strategie hanno dimostrato una notevole efficacia a dispetto della ridotta alterazione dei segnali in senso psicoacustico; il rovescio della medaglia è che la loro implementazione rimane relativamente più complessa rispetto alle più usate tecniche non lineari.

Nel bilancio tra semplicità realizzativa ed efficacia, è appunto la famiglia delle predistorsioni non lineari [40] che riveste un ruolo di primaria importanza. L'idea di base in questo caso è quella rappresentata in figura 2.6: ognuno dei due segnali x_i è separatamente sommato a una versione di sé stesso distorta non linearmente, attraverso un coefficiente di dosaggio α variabile a seconda dei casi. La distorsione non lineare è generalmente assunta diversa per i due rispettivi canali; un esempio, adottato anche nel presente lavoro, è dato da [41]

$$\tilde{x}_i = x_i + \alpha \cdot \frac{x_i - (-1)^i |x_i|}{2} \quad \text{per } i = 1, 2 \quad (2.18)$$

Variazioni sul tema sono state proposte in varie forme [40], comprendendo anche un metodo adattativo di selezione del parametro α [42], sulla base della

minimizzazione di una stima dinamica della funzione di coerenza.

E' riportato che le tecniche di pre-distorsione non lineare consentono di ridurre notevolmente il valore della funzione di coerenza, a dispetto della grande semplicità realizzativa. E' certamente intuibile, come del resto è stato sperimentalmente mostrato [42], che esiste una dipendenza diretta e monotona tra l'aumento del coefficiente di dosaggio α e la diminuzione del valore di $|\gamma(f)|$ mediato in frequenza. D'altra parte, valori eccessivamente elevati di α sono suscettibili di introdurre alterazioni tali da compromettere l'integrità percettiva dei segnali. Caso per caso, la scelta di α andrà effettuata in modo da non alterare eccessivamente la natura dei segnali da una parte, e dall'altra in modo da massimizzare il disaccoppiamento statistico.

Le vere implementazioni della SAEC non possono dunque prescindere dall'adottare un metodo artificiale di pre-elaborazione dei canali d'ingresso. Nonostante tali artifici consentano nella maggior parte dei casi di scongiurare la convergenza fittizia dei filtri, e garantiscano quindi, per $n \rightarrow \infty$, un'effettiva minimizzazione del disallineamento introdotto in (2.17), la matrice di correlazione stereofonica rimane generalmente ancora mal condizionata per quel che riguarda e influenza il tempo di convergenza degli algoritmi adattativi. La centralità della scelta di un algoritmo adattativo sufficientemente robusto rispetto al malcondizionamento di \mathbf{R} era chiara già dai primi contributi al problema [43]. In effetti se la necessità di una realizzazione in tempo reale è aggiunta all'insieme delle richieste, la SAEC richiede architetture estremamente robuste, spesso basate su algoritmi di tipo RLS e con strutture parallele a sottobande (quest'ultima utile perché mitiga, in ogni sottobanda, sia la divaricazione degli autovalori, che la complessità computazionale) [44, 45, 46].

E' bene chiarire che, anche a valle di tutte le tecniche di disaccoppiamento degli ingressi, l'onerosità computazionale e la complessità delle architetture necessarie alle realizzazioni pratiche è tuttora assai elevata. Soluzioni basate sulla semplificazione del problema, e su metodi di riduzione del problema al caso monofonico [47], seppur puntando ad obiettivi di cancellazione meno ambiziosi, possono facilmente essere preferiti proprio per la maggiore snellezza realizzativa. In questo panorama, l'uso della conoscenza del denominatore comune per l'ambiente di ricezione può introdurre semplificazioni architetture e guadagni di prestazioni. Il prossimo paragrafo dettaglierà questa possibilità, mostrando come proprio le difficoltà specifiche della SAEC qui ricordate ne fanno un'applicazione ideale per trarre il massimo del vantaggio dalla modellazione basata su CAP.

2.2.3 Una nuova architettura basata su CAP

I metodi di pre-eleborazione ricordati nel precedente paragrafo hanno tutti come obiettivo il disaccoppiamento statistico dei due segnali d'ingresso, x_1 e x_2 , all'interno del vincolo di tollerabilità dell'alterazione percettiva rispetto all'informazione spaziale codificata nei segnali stessi. Si è ricordato come il loro uso sia di fondamentale importanza per aggirare l'apparente impossibilità del problema della SAEC, ma continui a richiedere implementazioni adattative estremamente robuste e complesse, per arginare gli effetti del malcondizionamento di \mathbf{R} , matrice di correlazione stereofonica degli ingressi.

L'interesse sull'uso dei CAP in questo contesto nasce proprio dal problema centrale del malcondizionamento di \mathbf{R} . Più specificatamente, se a parità di L_W la dipendenza di $\text{cond}(\mathbf{R})$ da $|\gamma(f)|$ è stata in più sedi mostrata e dettagliata, l'abbattimento di $\text{cond}(\mathbf{R})$ basato sulla diminuzione della lunghezza L_W dei filtri adattativi è a tutt'oggi una stada non battuta in letteratura. La motivazione immediata di ciò è intrinseca in tutta la discussione iniziata nel paragrafo 2.2.2: se per i filtri \mathbf{w}_1 e \mathbf{w}_2 è assunta la forma FIR, come spiegato più volte una diminuzione di L_W conduce a una cancellazione d'eco virtuale, che porta il vettore \mathbf{w} ($2L_W \times 1$) lontano da \mathbf{h} ($2L_H \times 1$) e il dissallineamento in (2.17) lontano da zero.

Il discorso cambia se è ammessa una forma ricorsiva per le generiche strutture adattative $W_1(z)$ e $W_2(z)$. In particolare, usando la modellazione basata su CAP, e supponendo nota (e ovviamente costante) una stima $\hat{A}(z)$ del denominatore dei poli comuni $A(z)$, si può pensare a una realizzazione della SAEC come quella riportata in figura 2.7.

L'uso della stima del denominatore dei CAP è qui del tutto analoga a quella introdotta nel paragrafo 2.1.3 e in figura 2.4. Si ricorderà da lì che $C(z) = 1 - \hat{A}(z)$ ha innanzi tutto una comoda forma FIR da punto di vista realizzativo, e con quella topologia di rete non contribuisce a una maggiore e indesiderabile coloritura spettrale dei segnali d'ingresso.

Ripercorrendo i problemi associati alla SAEC illustrati nel paragrafo 2.2.2, emerge rapidamente che l'architettura appena proposta presenta significativi vantaggi, e in particolare:

- Dal punto di vista dell'algoritmo adattativo, promette di ridurre la lunghezza L_W , visto che in questo caso l' i -esimo filtro adattativo \mathbf{w}_i dovrà convergere non più verso l'intera risposta impulsiva di canale \mathbf{h}_i , ma verso

$$\mathbf{b}_i = [b_0^{(i)}, b_1^{(i)} \dots, b_{L_B}^{(i)}]$$

$$\text{dove} \quad b_0^{(i)} + b_1^{(i)} z^{-1} + \dots + b_{L_B}^{(i)} z^{-L_B} = B_i(z)$$

$$\text{e} \quad H_i(z) = B_i(z)/A(z), \quad \text{per } i = 1, 2$$

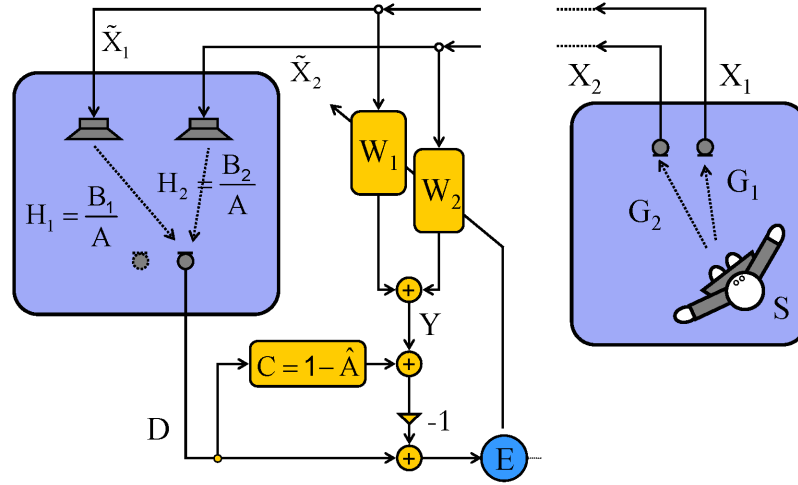


Figura 2.7: Realizzazione della SAEC basata sulla conoscenza di una stima $\hat{A}(z)$ del denominatore comune, $A(z)$, riportato esplicitamente nell'espressione delle RTF di canale $H_1(z)$ e $H_2(z)$. La dipendenza dalla variabile z è stata rimossa in tutti i simboli per maggiore chiarezza visiva.

ovvero il solo denominatore di canale, valendo chiaramente $L_B < L_H$ alla luce di quanto visto nel capitolo 1. In tal modo l'imposizione di $L_W \ll L_G$ risulta in parte naturale e giustificata.

- Dal punto di vista della capacità di modellazione delle risposte impulsive delle $H_i(z)$ con sufficiente dettaglio, in questo caso le lunghezze efficaci delle risposte impulsive del sistema cancellante non sono più date da L_W , ma dipenderanno anche e soprattutto dalla componente ricorsiva comune a cui contribuisce virtualmente $\hat{A}(z)$ in $C(z)$. La ridotta lunghezza dei \mathbf{w}_i , illustrata al primo punto, non ha quindi in questo caso ragione di indurre alcun effetto di convergenza verso soluzioni fittizie, nel senso dell'osservazione 2a del paragrafo 2.2.2.

Le idee appena esposte, insieme all'architettura di figura 2.7 costituiscono insieme uno dei contributi originali di questo lavoro di ricerca. Il prossimo paragrafo riporta le prove sperimentali usate per verificare l'effettiva efficacia di questo approccio. E' importante notare che tali prove avrebbero avuto una scarsa incidenza e un'altrettanto scarsa ripetibilità, se non avessero giovato dei risultati, anch'essi originali, esposti nel paragrafo 1.4 a proposito della selezione dell'ordine del denominatore dei CAP.

2.3 Set-up sperimentale e risultati

Come è stato appena ricordato, l'effettiva possibilità dell'uso della modellazione basata su CAP è debitrice del metodo di selezione dell'ordine introdotto nel paragrafo 1.4.2. Per coerenza con questa premessa, l'ambiente usato nelle simulazioni presentate parte dalla medesima configurazione di figura 1.2. Ciò permette di riusare tutti i risultati del paragrafo 1.4.3, sia rispetto alla selezione dell'ordine efficace che alle stime vere e proprie dei denominatori dei CAP.

Per quanto riguarda i canali H_1 e H_2 da modellare adattativamente, si è introdotto un nuovo microfono e si è selezionata una coppia di altoparlanti (rispettivamente 2 e 3 in figura 1.2), per rendere questi due canali-obiettivo esterni all'insieme originariamente usato per la ricerca dei CAP. Il nuovo set-up è illustrato in figura 2.8.

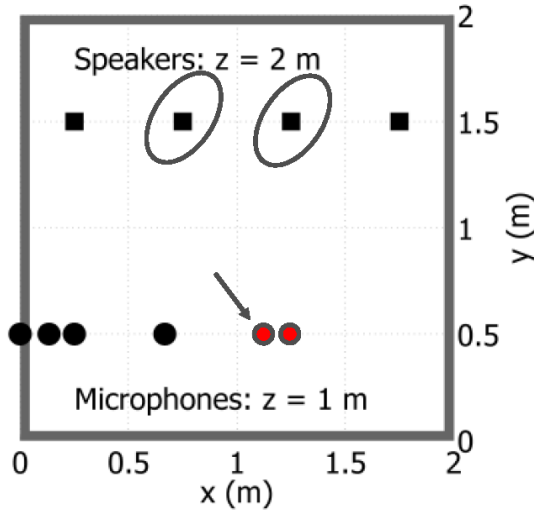


Figura 2.8: Set-up usato in simulazione come ambiente di ricezione, per l'architettura si SAEC proposta, basata su CAP. L'ambiente e la configurazione di base sono le stesse del paragrafo 1.4.3 e di figura 1.2. In più, è stata qui aggiunta una coppia microfonica. I 16 accoppiamenti originali sono quelli usati per la stima preventiva del denominatore comune, mentre il microfono indicato dalla freccia e gli altoparlanti cerchiati sono stati selezionati per ottenere i due rispettivi canali H_1 ed H_2 .

L'ambiente di trasmissione è stato simulato in modo analogo, con una nuova sorgente e la nuova coppia microfonica in figura 2.8.

Basandosi sui risultati del paragrafo 1.4.3, le simulazioni della SAEC sono state anch'esse ripetute per le cinque diverse larghezze di banda (250, 333, 500, 667, 1000) Hz. L'opportunità e l'efficacia della modellazione basata su CAP, come osservato nel paragrafo 1.4, ha infatti un'importante dipendenza dalla banda considerata,

ed è quindi importante che ogni architettura basata su CAP sia sottoposta a verifica al variare della larghezza di banda stessa.

L'architettura proposta riassunta in figura 2.7 è stata confrontata con la tradizionale implementazione non ricorsiva di figura 2.5. Perché tale confronto fosse praticabile, in tutte le simulazioni portate a termine è stata introdotta una predistorsione lineare degli ingressi, nel senso ricordato nel paragrafo 2.2.2, e più precisamente il metodo descritto dalla (2.18) con $\alpha = 0.5$.

Si ricorderà anche che le realizzazioni tradizionali tendono a usare strutture in sottobande a valle della pre-distorsione, perché la correzione del condizionamento di \mathbf{R} in (2.13) non è spesso sufficiente per ottenere tempi di convergenza ragionevoli con un singolo algoritmo a banda larga. Nelle prove qui presentate, finalizzate al puro confronto tra i metodi esistenti e la soluzione originale proposta, i criteri di primaria importanza sono stati la ripetibilità delle prove e la semplicità delle stesse.

Se la semplicità incoraggia all'uso di un singolo algoritmo a banda larga, l'ulteriore opportunità di ottenere tempi di convergenza gestibili dalle simulazioni ha portato alla scelta di rumore bianco per il segnale di sorgente S .

Un'altra importante variabile legata alle simulazioni effettuate è stata la scelta dell'algoritmo adattativo da usare per \mathbf{w} (cfr. equazione (2.11)). Nonostante la scelta del rumore bianco per la sorgente nell'ambiente di trasmissione, le criticità della SAEC, e in particolare il malcondizionamento di \mathbf{R} , impongono la scelta di algoritmi sufficientemente robusti.

D'altra parte, se lo scopo primario dell'introduzione dei CAP nella SAEC è proprio quello di abbassare il numero di condizionamento di \mathbf{R} , è interessante notare come tale effetto possa influenzare diversamente diversi algoritmi, con diverse caratteristiche di robustezza rispetto al malcondizionamento di \mathbf{R} .

In riferimento all'ultima osservazione fatta, sono stati scelti due algoritmi ampiamente noti e usati in letteratura: da una parte il già citato RLS e dall'altra l'APA (*Affine Projection Algorithm*), entrambi riadattati all'implementazione stereofonica [31, 41]. Come già ricordato, l'RLS è noto per eccellenza come uno degli algoritmi più robusti rispetto al malcondizionamento della matrice di correlazione degli ingressi; ciononostante, a causa dell'elevata complessità computazionale, è spesso usato in sue versioni semplificate (ad esempio il cosiddetto FRLS, o *Fast RLS* [41]). Implementazioni a complessità ridotta esistono anche per l'APA [28, 48, 49], che di per sé nasce come algoritmo a complessità variabile, a seconda dell'*ordine* usato, e che utilizza una stima dinamica di rango ridotto della matrice di correlazione. Tale caratteristica pone l'APA, sia dal punto di vista della complessità che da quello della robustezza, in una via di mezzo tra la versione *normalizzata* dell'LMS (o NLMS [23]) e l'RLS. L'ordine scelto, detto P e variabile tra 1 e L_W consente adattare il compromesso tra complessità e robustezza: $P = 1$ rende l'APA ridotto al semplice NLMS, mentre $P = L_W$

porta l'APA alla complessità dell'RLS.

Nel prove qui presentate, per semplificare la ripetibilità, entrambi gli algoritmi sono stati usati nelle forme originali; $P = 8$ è stato l'ordine selezionato per l'APA.

Per ogni larghezza di banda f_B e per ognuno dei due algoritmi adattativi, APA ed RLS, le prove sono state così organizzate:

1. E' stato simulato il sistema proposto in figura 2.7, usando per i \mathbf{w}_i la lunghezza L_B in tabella 1.1;
2. Per ogni simulazione è stato misurato l'andamento del disallineamento stereo $\mu_S[n]$ in (2.17) al variare dell'epoca di adattamento n , e stimato un valore $\mu_S[\infty]$, ovvero il valore stazionario (o minimo all'equilibrio) raggiunto da $\mu_S[n]$ ad adattamento avvenuto;
3. L'architettura tradizionale di figura 2.5 è stata simulata, scegliendo ora una lunghezza L_W tale da ottenere un disallineamento all'equilibrio $\mu_S[n]$ grossolanamente uguale a quello trovato nel caso precedente.
4. Il primo confronto tra le due simulazioni è stato basato sui due rispettivi valori di L_W usati nei due casi, detti rispettivamente L_{FIR} ed L_B ; tale confronto permette di mettere in luce il numero di coefficienti adattativi necessari nei due diversi casi per raggiungere la stessa accuratezza descrittiva, e permette di valutare due aspetti:
 - Quele sia il contributo del denominatore dei CAP nella riduzione della dimensionalità dei filtri, e quindi nella correzione del malcondizionamento della matrice di correlazione;
 - Quanto l'uso dei CAP consenta di ridurre la complessità dell'algoritmo adattativo: il peso computazionale di un algoritmo adattativo è sempre funzione della dimensione dei filtri, in modo strettamente dipendente dal particolare algoritmo scelto;
5. Un secondo confronto è stato incentrato sui tempi caratteristici di convergenza nei due diversi casi; ancora una volta, sulla scorta dei richiami fatti nel paragrafo 2.1.2, si ricorderà che il tempo di convergenza è funzione di $\text{cond}(\mathbf{R})$, e pertanto quest'analisi permette di valutare in modo diretto i benefici introdotti dall'uso dei CAP. I tempi caratteristici di convergenza possono essere valutati in diversi modi; qui si è scelto un semplice approccio, definito da:

$$n_\tau : \mu_S[n_\tau] = \mu_S[\infty] + 3 \text{ dB} \quad (2.19)$$

Il tempo caratteristico di convergenza è stato cioè misurato come quel valore dell'epoca di adattamento n_τ , per cui $\mu_S[n_\tau]$ è appena 3 dB superiore al valore di equilibrio $\mu_S[\infty]$.

La figura 2.9 presenta gli andamenti di $\mu_S[n]$ con l'uso dell'APA per alcuni diversi valori della banda f_B , mentre la figura 2.10 riporta le analoghe prove fatte con l'RLS. Per brevità, in entrambi i casi sono presentate le simulazioni per le sole tre rispettive larghezze di banda $f_B = 250, 500$ e 1000 Hz, mentre una sintesi completa dei risultati per tutte e cinque le bande è riportata nelle rispettive tabelle 2.1 per l'APA e 2.2 per l'RLS.

Tabella 2.1: Sintesi delle simulazioni della SAEC usando l'APA dell'8° ordine, con e senza il blocco C contenente il denominatore dei CAP.

f_B (Hz)	L_A	L_B	L_{FIR}	$\mu_S[\infty]$	n_τ ratio
250	28	18	49	-22 dB	4%
333	34	26	64	-23 dB	18%
500	49	40	91	-21 dB	19%
667	65	57	151	-33 dB	21%
1000	85	89	212	-31 dB	34%

L'analisi delle tabelle 2.1 e 2.2 mostra risultati di notevole interesse. Il primo aspetto a cui dedicare attenzione è la riduzione dei coefficiente adattativi necessari a raggiungere la stessa accuratezza descrittiva, evidente dal confronto tra la colonna L_{FIR} ed L_B separatamente per le due tabelle. Per tutte le larghezze di banda usate, la riduzione della lunghezze di L_W con l'uso dei CAP si aggira intorno al 60 – 65%, a significare che, con l'introduzione della stima del denominatore comune, la lunghezza dei filtri adattativi è pari a poco più di un terzo di quella originariamente necessaria.

Il secondo risultato rilevante riguarda i tempi caratteristici di convergenza, riportati nelle due tabelle in scala percentuale, come rapporto tra il tempo di convergenza n_τ con il metodo proposto e quello ottenuto con la struttura tradizionale, ovvero

$$n_\tau \text{ ratio} = \frac{n_\tau \text{ con CAP}}{n_\tau \text{ FIR}} [\%]$$

In questo caso è innanzi tutto importante notare come i risultati siano marcatamente diversi per i due rispettivi algoritmi adattativi. Nel caso dell'RLS, l'introduzione della conoscenza dei CAP non porta alcun eclatante beneficio in termini di tempo di convergenza, mentre nel caso dell'APA la riduzione va da circa 1/3 a circa 1/25, a seconda della banda. La differenza tra i due algoritmi

è una diretta conseguenza del differente grado di robustezza rispetto al condizionamento di \mathbf{R} : nel caso dell'RLS "puro", come in questo caso, si riescono a raggiungere tempi di convergenza accettabili anche con filtri lunghi, e la situazione non cambia un granché quando la riduzione della lunghezza dei filtri stessi abbassa $\text{cond}(\mathbf{R})$. Nel caso dell'APA, invece, la sensibilità rispetto a $\text{cond}(\mathbf{R})$ fa sì che una riduzione della lunghezza dei filtri si rifletta in maniera diretta sui tempi di convergenza.

Infine, sembra opportuno soffermarsi brevemente sull'andamento del tempo di convergenza al variare della banda f_B nel caso dell'APA. Si ricorderà che il problema della stima del denominatore dei CAP è impresa tanto più dura quanto maggiore è la larghezza di banda considerata, a causa della densità crescente con il quadrato della frequenza di banda f_B , descritta dalla (1.18); proprio da qui l'esigenza di un metodo per selezionare un ordine *efficace*, anche di molto inferiore a quello predetto della teoria. La differenza tra l'ordine teorico e quello efficace cresce naturalmente al crescere della frequenza. Prove sperimentali informali qui non riportate hanno mostrato che, usando il metodo HMK il rapporto tra il secondo e il primo ha un andamento grossolanamente proporzionale a f_B^2 . Tale effetto dipende dalla strutturale inadeguatezza di qualsiasi metodo di stima dei CAP attualmente disponibile. Questo può spiegare perché, al crescere di f_B , il guadagno in termini di tempo di convergenza sia, seppur sempre importante, progressivamente sempre più contenuto.

2.4 Conclusioni

Globalmente, i risultati presentati mostrano che l'introduzione dei CAP nella SAEC permette di conseguire in modo congiunto due importanti risultati:

- Ridurre la lunghezza dei filtri adattativi a parità di accuratezza descrittiva;

Tabella 2.2: Sintesi delle simulazioni della SAEC usando l'RLS, con e senza il blocco C contenente il denominatore dei CAP.

f_B (Hz)	L_A	L_B	L_{FIR}	$\mu_S[\infty]$	n_τ ratio
250	28	18	56	-26 dB	62%
333	36	24	74	-24.5	87%
500	49	40	96	-25 dB	81%
667	65	57	160	-33.5 dB	90%
1000	85	89	220	-32 dB	89%

- Diminuire i tempi caratteristici di convergenza

Per i motivi ricordati, il secondo effetto è meno evidente per gli algoritmi adattativi più robusti rispetto al condizionamento di \mathbf{R} , che finiscono però normalmente per essere usati con difficoltà, a causa dell'elevata complessità computazionale. Il costo realizzativo pagato per l'introduzione di tali miglioramenti è da una parte quello della stima del denominatore dei CAP (Portata a termine una volta per tutte, e senza incidenza nell'elaborazione in tempo reale), e dall'altra quella dell'introduzione di un filtro FIR statico ($C(z)$ in figura 2.7), comune ai due canali stimati, di poco conto nel bilancio computazionale e dalla certa stabilità. Come già ricordato, una limitazione del metodo proposto, a sua volta funzione degli esistenti metodi di stima dei CAP, è la sempre minore incidenza per larghezze di banda crescenti. Se da una parte tale limitazione verrebbe parzialmente rimossa dall'invenzione di metodi di stima più robusti, dall'altra un grande numero di implementazioni di SAEC è già implementato a doppia o a multipla banda, essendo la parte a bassa frequenza spesso riconosciuta come la più problematica e la principale candidata, anche per ragioni percettive, ad essere trattata in modo strettamente stereofonico.

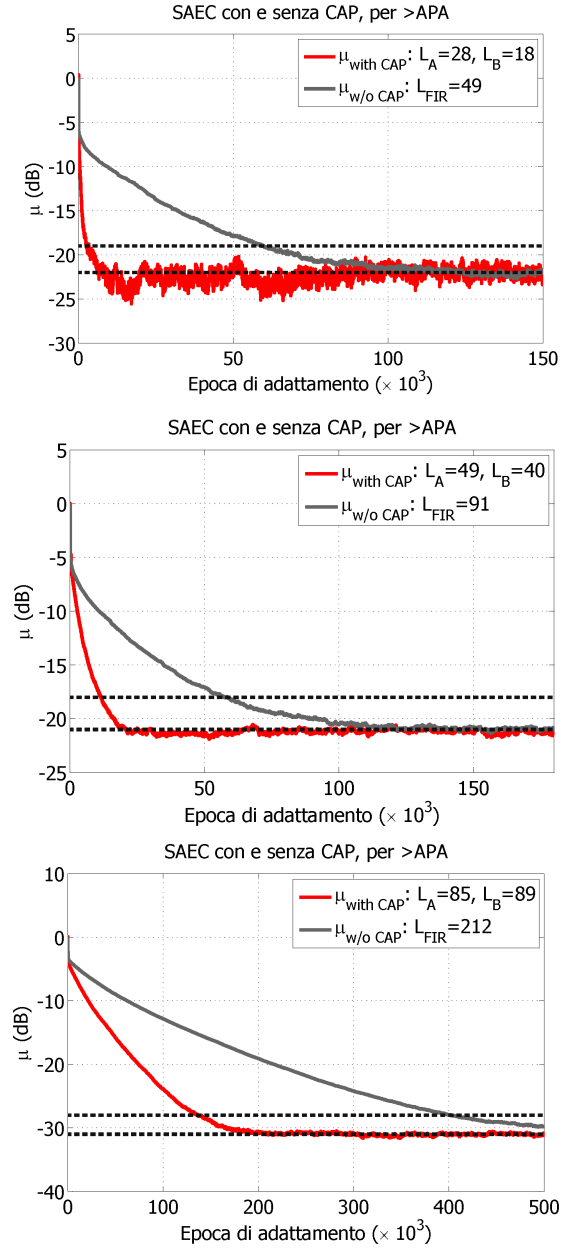


Figura 2.9: Andamenti di $\mu_S[n]$ con l'uso dell'APA di ordine 8, per le tre diverse larghezze di banda $f_B = 250$ Hz (in alto), $f_B = 500$ Hz (al centro) e $f_B = 1$ kHz (in basso); in tutti e tre i casi l'andamento rosso è quello della SAEC basata su CAP, mentre il riferimento in nero è l'andamento dell'architettura tradizionale. Le linee tratteggiate mostrano i due livelli $\mu_S[n_\tau]$ e $\mu_S[\infty]$, rispettivamente.

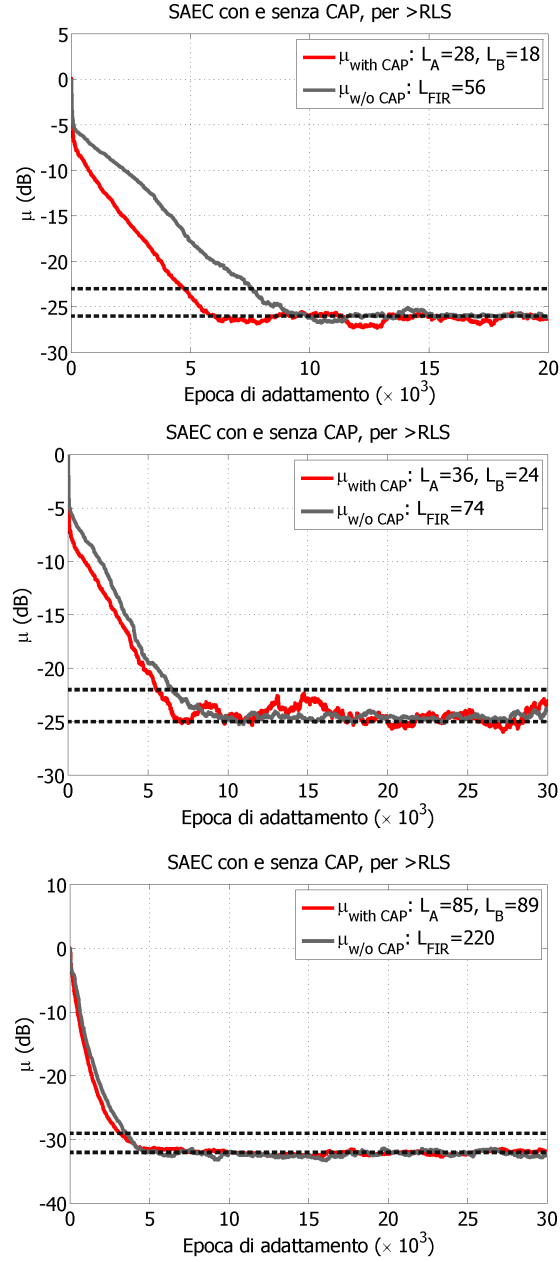


Figura 2.10: Andamenti di $\mu_S[n]$ con l'uso dell'RLS, per le tre diverse larghezze di banda $f_B = 250$ Hz (in alto), $f_B = 500$ Hz (al centro) e $f_B = 1$ kHz (in basso); in tutti e tre i casi l'andamento rosso è quello della SAEC basata su CAP, mentre il riferimento in nero è l'andamento dell'architettura tradizionale. Le linee tratteggiate mostrano i due livelli $\mu_S[n_\tau]$ e $\mu_S[\infty]$, rispettivamente.

Capitolo 3

Verso un modello analitico basato sui Poli Acustici Comuni

3.1 Introduzione

Il primo capitolo di questo lavoro, e in particolare il criterio di selezione dell'ordine nella stima di CAP, ha lasciato ancora aperti numerosi interrogativi relativi all'opportunità e alla liceità della modellazione di sistemi acustici multicanale basata su CAP.

I principali problemi aperti derivano dall'assenza di soluzioni analitiche per l'equazione di Helmholtz (1.13) e dall'esser vincolati a far riferimento al corrispondente problema agli autovalori, a sua volta dipendente dalla geometria dell'ambiente considerato. Con particolare riferimento ai temi della non parsimonia e della non disponibilità di un'espressione per i numeratori (cfr. paragrafo 1.3.6), l'assenza di una formulazione analitica per la soluzione dell'equazione di Helmholtz sembra costituire il fattore con maggiore incidenza.

In questo capitolo sono presentate delle idee originali, legate proprio alla possibilità di costruire un'espressione analitica della soluzione dell'equazione delle onde per sistemi semplici. Per tale espressione, definita costruttivamente, e non per risoluzione diretta della PDE e delle relative condizioni al contorno, è presentata una naturale estensione della soluzione del corrispondente problema unidimensionale, anch'essa derivata in questo contesto come contributo originale.

I risultati presentati in questo capitolo forniscono quindi da un lato delle parziali risposte agli interrogativi rimasti aperti nel capitolo 1, e dall'altro offrono un banco di prova significativamente innovativo per la valutazione dell'efficacia dei metodi di stima di CAP in generale, come verrà chiarito nel capitolo 4.

3.2 Poli Comuni per sistemi 1-D

Come punto di partenza di questo capitolo, si consideri il semplice sistema acustico unidimensionale descritto dallo schema di figura 3.1. Il tubo, scelto come

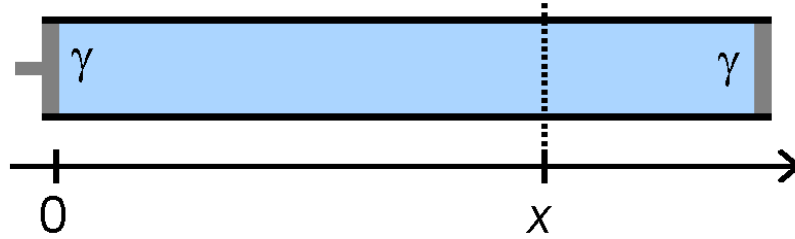


Figura 3.1: Un semplice sistema unidimensionale: tubo con sorgente (pistone mobile) in un estremo, e terminazioni parzialmente riflettenti secondo il coefficiente di riflessione in ampiezza γ .

particolare riduzione al caso 1-D del problema della stanza parallelepipedale in 3-D, è supposto di lunghezza ℓ , pieno d'aria, e con estremità né totalmente assorbenti né totalmente riflettenti. Scelto un sistema di riferimento unidimensionale con origine nell'estremo sinistro, e dette $p^+(t, x)$ e $p^-(t, x)$ rispettivamente le onde di pressione progressiva e regressiva, si definisca il coefficiente di riflessione in ampiezza γ attraverso la seguente espressione

$$\gamma = \frac{p^+(t, 0)}{p^-(t, 0)} = \frac{p^-(t, \ell)}{p^+(t, \ell)} < 1 \quad (3.1)$$

Per semplicità di trattazione, si immagina che la sorgente acustica sia posta in un estremo: la parete nell'origine si assumerà mobile e manovrata da un pistone, come rappresentato in figura.

A partire da questo semplice sistema a sorgente unica, si intende definire per costruzione una generica risposta impulsiva nel dominio s , ottenuta posizionando un sensore di pressione in una qualsiasi posizione x .

Per ispezione, se si immagina un impulso unitario partire al tempo $t = 0$ dalla sorgente nell'origine, i primi contributi nel tempo registrati dal sensore in

x saranno, nell'ordine, i seguenti:

$$\begin{aligned} P_0(s; x) &= e^{-s\tau_x} \\ P_1(s; x) &= \gamma e^{-s\tau_{2\ell-x}} \\ P_2(s; x) &= \gamma^2 e^{-s\tau_{2\ell+x}} \\ P_3(s; x) &= \gamma^3 e^{-s\tau_{4\ell-x}} \\ &\dots \end{aligned}$$

Nelle precedenti espressioni si è adottata la seguente convenzione per esprimere il tempo necessario a coprire la generica lunghezza λ

$$\tau_\lambda = \frac{\lambda}{c}$$

dove $c \simeq 340 \text{ ms}^{-1}$ rappresenta come al solito la velocità del suono in aria.

Un raggruppamento dei termini di ordine rispettivamente pari e dispari conduce alla seguente serie

$$\begin{aligned} H(s; x) &= \sum_{j=0}^{\infty} P_j(s; x) \\ &= \sum_{i=0}^{\infty} \left(\gamma^{2i} e^{-s \left[\frac{2Li+x}{c} \right]} + \gamma^{2i+1} e^{-s \left[\frac{2L(i+1)-x}{c} \right]} \right) \end{aligned} \quad (3.2)$$

La (3.2) può risciversi dopo alcuni passaggi come

$$H(s; x) = \sum_{i=0}^{\infty} e^{-s \frac{x}{c}} \left[\gamma^2 e^{-s \frac{2\ell}{c}} \right]^i + \sum_{i=0}^{\infty} e^{-s \frac{2\ell-x}{c}} \left[\gamma^2 e^{-s \frac{2\ell}{c}} \right]^i \quad (3.3)$$

Portando fuori dalla sommatoria i termini indipendenti da i , mettendo in evidenza la comune sommatoria e ricordando che

$$\left| \gamma^2 e^{-s \frac{2\ell}{c}} \right| < 1 \quad \Rightarrow \quad \sum_{i=0}^{\infty} \left[\gamma^2 e^{-s \frac{2\ell}{c}} \right]^i = \frac{1}{1 - \gamma^2 e^{-s \frac{2\ell}{c}}}$$

si ottiene infine

$$H(s; x) = \frac{e^{-s\tau_x} + \gamma e^{-s(\tau_{2\ell}-\tau_x)}}{1 - \gamma^2 e^{-s\tau_{2\ell}}} \quad (3.4)$$

La (3.4) rappresenta la soluzione analitica per la risposta impulsiva nel dominio s (quindi nient'altro che la funzione di trasferimento acustica, o RTF), per il canale costituito dall'accoppiamento tra la sorgente in $x = 0$ e un generico sensore nella posizione x .

La prima osservazione degna di nota, a conferma dei ragionamenti del paragrafo 1.3, è che ci si trova in presenza di una soluzione trascendente, e non razionale.

Se normalmente, per “funzioni di rete” in reti a costanti concentrate (RCC), il passaggio al dominio della trasformata z avviene preferenzialmente tramite

l'uso di trasformazioni lineari o bilineari (che conservano la forma razionale fratta e il relativo ordine), in questo caso è proprio mantenendo la definizione trascendente di $z = e^{sT_C}$ che si riesce a ricondurre la (3.4) ad una forma razionale in z . Assumendo infatti

$$\frac{\tau_x}{T_C} = N_x \quad \text{e} \quad \frac{\tau_{2\ell}}{T_C} = N_{2\ell} \quad \text{con} \quad N_x, N_{2\ell} \in \mathbb{N} \quad (3.5)$$

la (3.4) può trasformarsi nella conveniente forma

$$H(z; x) = \frac{z^{-N_x} + \gamma z^{-(N_{2\ell} - N_x)}}{1 - \gamma^2 z^{-N_{2\ell}}} \quad (3.6)$$

La (3.6) è di grande importanza per diversi motivi, tra i quali:

1. Fornisce, come anticipato, un'espressione analitica della funzione di trasferimento del sistema per una qualsiasi posizione x del sensore, in forma razionale in z ;
2. Contiene un denominatore completamente indipendente dalla posizione del sensore, e quindi dal particolare canale acustico scelto per il sistema considerato, supportando l'ipotesi della liceità della modellazione basata su CAP e offrendone anzi un'espressione analitica;
3. Offre un'espressione della generica RTF che, oltre a mostrare analiticamente i CAP, è massimamente parsimoniosa, perché definisce il sistema in oggetto con i due soli parametri γ ed $N_{2\ell}$;
4. Restituisce un'espressione per il numeratore del singolo canale, dalla chiara interpretazione intuitiva.

Alcune delle tematiche racchiuse nei punti appena elencati verranno sviluppati in dettaglio nel paragrafo che segue, dedicato alla deduzione e all'esplorazione di tutte le indicazioni quantitative contenute nella (3.6).

3.3 Analisi delle radici della soluzione razionale 1-D

La prima facile considerazione legata alla (3.6) deriva dall'osservazione del numeratore, $B(z; x)$. A dispetto di ogni informazione resa disponibile dalle normali trattazioni del problema 3-D, questa espressione ne rivela una semplice forma, e più precisamente l'essere composto da due termini:

- Un primo termine, z^{-N_x} , corrispondente al fronte sonoro impegnato nel percorso diretto;
- Un secondo termine, $\gamma z^{-(N_{2\ell} - N_x)}$, per il percorso corrispondente alla prima riflessione

Il resto dei fenomeni, dall'arrivo del fronte proveniente dalla seconda riflessione in poi, è descritto in termini puramente ricorsivi, ovvero attraverso la moltiplicazione di questo semplice numeratore per $1/A(z)$, dove $A(z)$ è proprio il denominatore comune del sistema. Questa prima osservazione sarà molto importante per le considerazioni che seguiranno nel paragrafo 3.4.

Sempre a proposito del numeratore, appare importante notare che l'ordine, pari a $N_{2\ell} - N_x$, dipende dal particolare canale e può essere al più pari all'ordine del denominatore, nel caso in cui il sensore è sovrapposto alla sorgente in $x = 0$.

Un altro importante passo consiste nel derivare le proprietà delle radici del denominatore comune, rappresentanti proprio i CAP del particolare sistema considerato. Per trovare i poli basta naturalmente imporre

$$A(z) = 1 - \gamma^2 z^{-N_{2\ell}} = 0 \quad \Leftrightarrow \quad z_{\pm k}^{-N_{2\ell}} = \frac{1}{\gamma^2} e^{\pm j2\pi k}$$

Questo porta, per la k -esima coppia di soluzioni complesse coniugate $z_{\pm k}$ all'espressione

$$z_{\pm k} = \gamma^{\frac{2}{N_{2\ell}}} \cdot e^{\mp j2\pi \frac{k}{N_{2\ell}}} \quad (3.7)$$

riportata graficamente in figura 3.2. Si noti che i poli trovati sono in numero $N_{2\ell} = F_C \cdot \tau_{2\ell}$, intero per le ipotesi fatte e dipendente linearmente dalla frequenza di campionamento F_C .

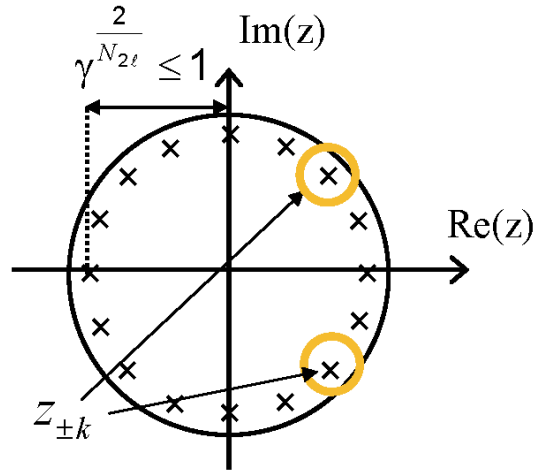


Figura 3.2: Le posizioni dei CAP del sistema 1-D nel piano z , così come descritte dalla (3.9).

A sua volta la (3.7), esplicitata l'espressione di $N_{2\ell}$ in funzione di T_C e $\tau_{2\ell}$, prende la forma intermedia

$$z_{\pm k} = \exp \left[T_C \left(\frac{2}{\tau_{2\ell}} \log(\gamma) \mp j \cdot \frac{2\pi k}{\tau_{2\ell}} \right) \right] \quad (3.8)$$

Con la finalità di rendere più intuitiva la lettura delle posizioni delle coppie di poli trovate, la (3.8) permette di tornare agevolmente al dominio continuo s

$$s_{\pm k} = \frac{1}{\tau_{2\ell}} (2 \log(\gamma) \mp j \cdot 2\pi k) \quad (3.9)$$

La prima conseguenza dell'ultima espressione trovata è la semplicità di rappresentazione nel piano s . Le coppie in (3.9) sono rappresentate in figura 3.3.

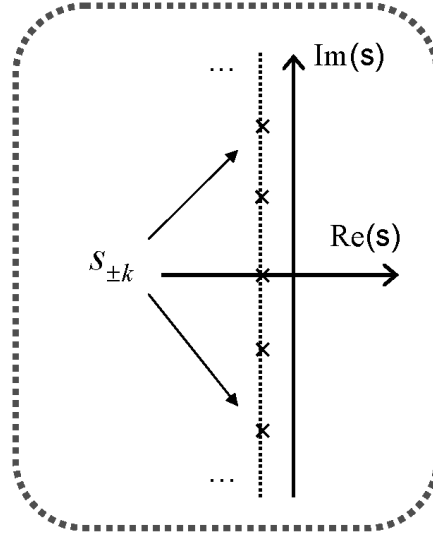


Figura 3.3: Le posizioni dei CAP del sistema 1-D nel piano s , così come descritte dalla (3.9).

Il modulo della k -esima coppia è approssimato sempre meglio dalla parte immaginaria al crescere di k , tanto da permettere di scrivere

$$\Omega_k = |s_{\pm k}| \approx \text{Im}(s_{\pm k}) = 2\pi \frac{k}{\tau_{2\ell}} = 2\pi \cdot \frac{c}{2} \cdot \frac{k}{\ell} \quad (3.10)$$

Per quanto riguarda il k -esimo fattore di merito Q_k , quest'ultimo è facilmente calcolabile nel dominio s , ed è dato da

$$Q_k = \frac{|s_{\pm k}|}{-2 \text{Re}(s_{\pm k})}$$

Sostituendo nella precedente espressione la (3.9) si arriva facilmente a

$$Q_k = \frac{\sqrt{(\log \gamma)^2 + k^2 \pi^2}}{|\log \gamma|} = \sqrt{1 + \left(\frac{k\pi}{\log \gamma}\right)^2} \quad (3.11)$$

Ipotizzando a questo punto di trovarsi in condizioni non eccessivamente dissipative si potrà assumere

$$\gamma \approx 1 \quad \Rightarrow \quad \log \gamma \approx \gamma - 1$$

e, infine

$$Q_k = \sqrt{1 + \left(\frac{k\pi}{\gamma - 1}\right)^2} \approx \frac{k\pi}{1 - \gamma} \quad (3.12)$$

La (3.12) e la (3.10) costituiscono insieme un punto d'arrivo molto importante per questo paragrafo. Mentre la (3.10) non fa altro che confermare conoscenze già disponibili (basta infatti considerare la (1.16) per $k_2 = k_3 = 0$ per ritrovare la stessa identica espressione per $F_k = \Omega_k/(2\pi)$), la (3.12) permette di formulare un'osservazione nuova: il fattore di merito dei poli comuni in un sistema unidimensionale cresce linearmente con il numero d'onda k . L'importanza del confronto con la (1.16), cionondimeno, risiede nell'indicazione che possano ragionevolmente estendersi ad ambienti 3-D risultati dimostrati analiticamente solo in semplici sistemi 1-D.

Prima di procedere oltre sembra importante chiarire che le osservazioni contenute in questo paragrafo sono integralmente frutto dell'attività legata a questo lavoro; questa è tra l'altro la ragione dell'esiguo numero di citazioni incontrate fin qui. Pur non potendo affermare che gran parte delle espressioni analitiche in sé siano del tutto nuove (trattazioni di questo tipo possono rintracciarsi in diversi ambiti disciplinari: basti pensare al caso delle linee di trasmissione elettriche [16]), si desidera precisare che le considerazioni da lì derivate e la contestualizzazione all'ambito della modellazione basata su CAP sono originali e innovative, e a parere di chi scrive introducono un gran contributo di chiarezza nella giustificazione della modellazione CAP, come appare chiaro dalle risposte fornite già fin qui ai maggiori interrogativi lasciati aperti nel paragrafo 1.3.

Il paragrafo seguente guiderà il lettore in un altro passo utile a chiarire ancora meglio come le deduzioni raggiunte fin qui possano aiutare la modellazione basata su CAP di sistemi acustici ambientali tridimensionali.

3.4 Estensione ad ambienti parallelepipedali dissipativi in 3-D

Nel paragrafo precedente, a partire dalle ipotesi (3.5) e per il semplice sistema unidimensionale in figura 3.1 si è mostrato che:

- E' possibile costruire una soluzione per l'equazione delle onde che restituisce funzioni di trasferimento razionali nel dominio z
- Che tale soluzione ha un denominatore indipendente dal particolare *canale* scelto, ed è pertanto coerente con la modellazione CAP, con un ordine dipendente (qui linearmente) dalla larghezza di banda f_B (ovvero dalla frequenza di campionamento F_C)
- Che le radici del denominatore così trovato, descritte nel dominio s alle espressioni (3.10) e (3.12), sono in completo accordo con le informazioni disponibili dalla risoluzione del problema agli autovalori dell'equazione di Helmholtz, nel caso di ambiente parallelepipedale

L'ultimo punto, in particolare, costituisce il tema di interesse centrale di questo paragrafo. Ripercorrendo la sintesi fatta nel paragrafo 1.3, e in particolare in 1.3.4, si ricorderà che il caso di ambiente non dissipativo di forma parallelepipedale costituisce uno dei pochi sistemi per cui il problema agli autovalori dell'equazione di Helmholtz è facilmente risolvibile. In quel caso l'ammissione preventiva di assenza di perdite alle pareti portava a un'espressione per le infinite frequenze modali in funzione dei tre indici interi k_1, k_2, k_3 , che si riporta qui per comodità:

$$F_{k_1, k_2, k_3} = \frac{c}{2} \sqrt{\left(\frac{k_1}{L_1}\right)^2 + \left(\frac{k_2}{L_2}\right)^2 + \left(\frac{k_3}{L_3}\right)^2} \quad k_i = 0 \dots \infty \quad (3.13)$$

Considerando la soluzione del problema unidimensionale come caso particolare della (3.13) con $k_2 = k_3 = 0$ e $k_1 = k$, è evidente che quest'ultima è identica alla (3.10) divisa per 2π in entrambi i membri, ovvero:

$$F_k \approx \frac{k}{\tau_{2\ell}} = \frac{c}{2} \cdot \frac{k}{\ell} \quad (3.14)$$

Le ragioni per cui la (3.14) può meritare di rivestire più interesse dell'apparentemente più generica e nota (3.13) sono almeno due:

- La (3.14) è stata ottenuta per costruzione, senza alcuna ipotesi di assenza di perdite, pur descrivendo, in questa forma, l'approssimazione per basse perdite (o alto k);
- La (3.9) è derivata dall'espressione analitica del denominatore della (3.6): in tal modo porta con sé la conoscenza non solo delle frequenze modali, ma anche dei poli veri e propri, e quindi dei fattori di merito in (3.11), approssimati dalla (3.12) come

$$Q_k \approx \frac{k\pi}{1 - \gamma} \quad (3.15)$$

sempre per basse perdite o per alti valori di k .

Dal confronto tra la (3.13) e la (3.14) nasce un'indicazione per l'estensione della (3.15) al caso 3-D. Il fatto che Q_k cresca proporzionalmente al numero d'onda k , in particolare, porta con sé un ben più importante osservazione candidata ad essere assunta invariante al crescere della dimensionalità del problema. la proporzionalità tra Q_k ed F_k . Sulla base di quest'assunzione, pur non giustificata qui dal punto di vista analitico nasce la ragionevole estensione al caso 3-D per il fattore di merito:

$$Q_{k_1, k_2, k_3} = \pi \sqrt{\left(\frac{k_1}{1 - \gamma_1}\right)^2 + \left(\frac{k_2}{1 - \gamma_2}\right)^2 + \left(\frac{k_3}{1 - \gamma_3}\right)^2} \quad (3.16)$$

L'altra caratteristica identificata nella soluzione analitica del problema 1-D costruita in questo capitolo è la forma del denominatore della (3.6). Si è osservato in precedenza come i due termini del numeratore abbiano una facile interpretazione intuitiva: il primo rappresenta il fronte portato dal cammino propagativo diretto, mentre il secondo aggiunge il contributo portato dal fronte della cosiddetta *prima riflessione*.

In tre dimensioni tale forma è facilmente estendibile, con le seguenti precisazioni:

- E' naturale richiedere il riferimento a un modello di ambiente 3-D che rappresenti la naturale estensione del tubo con un pistone in un estremo di figura 3.1. L'analogo nel caso parallelepipedale è rappresentato in figura 3.4, e prevede una sorgente in un vertice come estensione del pistone in un estremo del tubo.
- Il nuovo sistema in 3-D mantiene naturalmente un singolo contributo portato da un cammino di propagazione diretto (privo di riflessioni), ma richiede l'introduzione di tre contributi di prima riflessione, relativi rispettivamente ad ognuna delle tre pareti non adiacenti al vertice in cui si trova la sorgente.
- Per ognuno dei tre termini di prima riflessione si potrà introdurre un diverso coefficiente γ_i ($i = 1, 2, 3$)
- Nel caso 3-D è importante tener conto dell'attenuazione proporzionale all'inverso del quadrato della distanza in tutti e quattro i termini, dovuta alla forma sferica, e non più piana, dei fronti d'onda.

In sintesi, il numeratore della RTF di canale prenderà la forma seguente

$$B(z; \mathbf{r}) = z^{-N_r} + \sum_{i=1}^3 \alpha_i(\mathbf{r}, \mathbf{L}) \gamma_i z^{-N_i(\mathbf{r})} \quad (3.17)$$

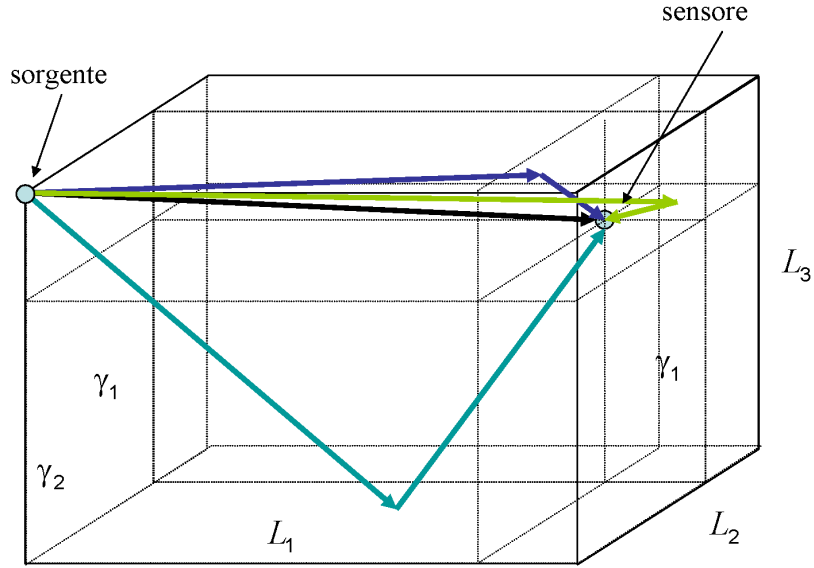


Figura 3.4: Un semplice sistema 3-D che rappresenti una fedele estensione del semplice sistema 1-D di figura 3.1.

con $N_r = F_C \cdot r/c$, $r = \|\mathbf{r}\|$, dove $\mathbf{r} = (x, y, z)$ rappresenta la posizione del sensore, se l'origine è supposta coincidente con la posizione della sorgente. $N_i(\mathbf{r})$ rappresenta il numero di campioni corrispondente al ritardo associato con l' i -esimo cammino di prima riflessione (supposto per semplicità intero), γ_i il coefficiente di riflessione relativo all' i -esima coppia di pareti, e $\alpha_i(\mathbf{r}, \mathbf{L})$ il termine di attenuazione quadratica associato all' i -esimo percorso di prima riflessione, funzione della posizione del sensore e delle dimensioni lineari dell'ambiente, $\mathbf{L} = (L_1, L_2, L_3)$.

La (3.17), insieme alle considerazioni che hanno condotto alla (3.16), permettono dunque l'estensione della forma analitica per le RTF di canale date dalla (3.6) al modello di ambiente confinato parallelepipedale di figura 1.1. Questo risultato rappresenta un passo di grande importanza per diversi motivi. Innanzi tutto, benché nel corso di questo capitolo si siano fatte numerose ipotesi semplificative, entro questa limitazione si può finalmente affermare di aver dato una risposta a tutti gli interrogativi teorici rimasti aperti dal paragrafo 1.3. Oltre a quanto è stato già notato in precedenza, si noti che, sulla base della (3.17) e della (3.6), è possibile dare un'interpretazione fisica dell'ordine del numeratore per le RTF basate su CAP, e fornirne una relativa quantificazione. La massima lunghezza di un percorso di prima riflessione in 1-D è infatti *al più* pari al numero di CAP (ovvero all'ordine del denominatore comune).

Quando la dimensionalità del problema aumenta, il più lungo cammino di prima riflessione continua ad avere una dipendenza dalle dimensioni lineari del sistema (ad esempio $\sqrt{(4L_1^2 + L_2^2)}$ nel caso 2-D, per $L_1 > L_2$), mentre l'ordine

del denominatore CAP è proporzionale alla misura di ordine massimo (ovvero la superficie in 2-D e il volume in 3-D), come del resto già indicato dalla (1.17).

Per un ambiente 3-D parallelepipedale con lunghezze $L_1 > L_2 > L_3$ l'ordine del numeratore si potrà considerare sempre ragionevolmente limitato da

$$\max(L_B) = \frac{F_C}{c} \sqrt{4L_1^2 + 4L_2^2 + L_3^2} \quad (3.18)$$

Oltre alla luce fatta sull'ordine dei numeratori, le espressioni trovate racchiudono un'altra fondamentale importanza: per la prima volta esse permettono di generare dei test per l'analisi delle prestazioni dei metodi di stima dei denominatori comuni, da una parte conoscendo il denominatore da stimare, e dall'altro controllando coerentemente i singoli numeratori di canale. E' infatti di fondamentale importanza notare che provare l'efficacia di un metodo di stima CAP con numeratori di canale dalla forma non realistica può alterare completamente il processo di validazione.

Questo aspetto sarà usato nell'ultimo capitolo di questo lavoro, che, tra le altre cose, è anche dedicato ad investigare sull'effettiva accuratezza del metodo HMK, considerato in letteratura il metodo di riferimento per la stima di CAP.

E' infine da ricordare che numerose delle considerazioni fatte negli ultimi paragrafi si riferiscono all'interpretazione teorica delle RTF acustiche basate su CAP, e non devono rischiare di generare confusione con l'uso pratico della modellazione basata su CAP, specialmente riguardo alla scelta dell'ordine. E' importante non dimenticare che l'ordine efficace definito nel capitolo 1 è sempre notevolmente inferiore all'ordine teorico.

Capitolo 4

Metodi di stima del denominatore comune

4.1 Problemi aperti nella stima di poli comuni

Più volte, nei precedenti capitoli, sono stati citati problemi legati alle limitazioni imposte dai metodi di stima di CAP attualmente rintracciabili in letteratura. In particolare, le prove condotte con l'uso del metodo HMK, sia per quanto ha riguardato il problema della scelta dell'ordine (Cap. 1), che nella realizzazione di un'architettura alternativa per la cancellazione d'eco stereofonica (Cap. 2), hanno messo in luce limitazioni raggruppabili nelle seguenti quattro aree:

1. La densità modale lungo l'asse della frequenza, per sistemi acustici distribuiti e a tre dimensioni spaziali, è un fattore di criticità nella stima delle stesse frequenze modali, da due punti di vista
 - In un modo già noto alle tecniche di stima spettrale parametrica [19], compromette la risoluzione della stima delle coppie di poli complessi coniugati associate alle frequenze modali, o ai coefficienti del denominatore di cui tali poli sono le radici. Particolarmente, questo limita l'ordine del denominatore dei CAP stimabile, obbligando a puntare alla selezione di un ordine *efficace* notevolmente inferiore all'ordine reale;
 - Richiede la stima di un numero di parametri liberi (i coefficienti del denominatore comune nel caso del metodo HMK) comunque molto elevato, ponendo potenziali problemi dal punto di vista della realizzabilità computazionale degli stessi metodi di stima.

In particolare rispetto al secondo punto, si ricorderà che il metodo HMK si basa sulla pseudo-inversione di una matrice di dimensioni assai elevate,

citare in (1.8) e qui riportate per comodità

$$N_H (L_A + L_H + 1) \times [L_A + N_H (L_B + 1)] \quad (4.1)$$

Proprio il consumo di risorse computazionali legato a tale operazione confina l'applicabilità del metodo entro i limiti dei problemi di dimensionalità (ovvero di ordini L_A ed L_B) ridotta, e quindi, conseguentemente, con ridotta larghezza di banda. Una prima esigenza che stimola la ricerca di nuovi metodi di stima dei CAP è quella legata al potenziale uso di ordini più elevati.

2. Oltre al problema degli ordini elevati, le dimensioni della matrice usata dal metodo HMK in (4.1) limitano anche la scelta dell'ambiente acustico considerato dal punto di vista delle proprietà riverberanti. Ambienti con pareti non attenuanti rischiano infatti di produrre risposte impulsive di lunghezza efficace L_H tutt'altro che contenuta, con conseguente insostenibilità delle dimensioni in (4.1). Una seconda esigenza che motiva la ricerca di nuovi metodi è quindi la possibilità di usare risposte impulsive di lunghezze realistiche, per poter usare modelli basati su CAP in sistemi acustici legati ad ambienti a loro volta realistici.
3. Considerando un sistema come quello di figura 4.1 il metodo HMK si fonda sulla conoscenza preventiva di un certo numero di canali acustici nell'ambiente scelto, nella forma di risposte impulsive $h_i[k]$, per $k = 0, \dots, L_H - 1$ e $i = 1, \dots, N_H$, essendo N_H il numero dei canali-campione disponibili. La stima di risposte impulsive è una procedura preventiva che, seppur non limitante dal punto di vista teorico, può richiedere l'uso di risorse computazionali e infrastrutturali aggiuntive. Una caratteristica interessante di nuovi metodi di stima potrebbe essere la capacità di usare semplicemente i segnali di uscita dei canali-campione, $x_O^{(i)}[n]$ per $i = 1, \dots, N_H$ potenzialmente anche ammettendo di conoscere i relativi segnali d'ingresso (Qui unicamente $x_I[n]$, essendo per semplicità il sistema in figura 4.1 di tipo SIMO, ovvero a singolo ingresso e più uscite).
4. Come naturale estensione delle considerazioni al precedente punto, e sempre in riferimento allo schema di figura 4.1, un'altra caratteristica vantaggiosa sarebbe quella di poter sfruttare i segnali di uscita dei canali, $x_O^{(i)}[n]$ con metodi di tipo *Blind* (o non supervisionato), ovvero escludendo la necessità della conoscenza del segnale d'ingresso $x_I[n]$. Ciò potrebbe introdurre notevoli semplificazioni nella realizzabilità pratica della stima di CAP.

I punti appena elencati si riferiscono da una parte a problemi associati specificatamente al metodo HMK, ma dall'altra presentano idee desiderabili in generale per potenziali nuovi metodi di stima.

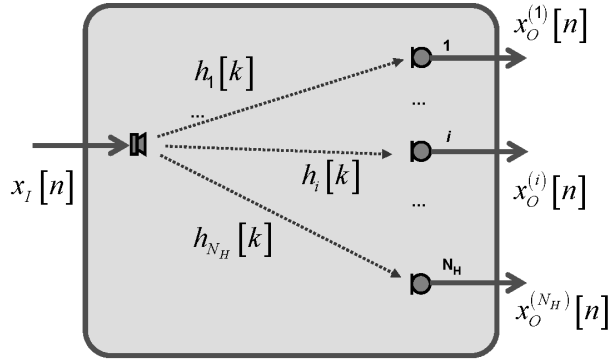


Figura 4.1: Un esempio di semplice sistema SIMO usabile per la definizione di N_H canali in un ambiente assegnato. Sono evidenziati il segnale d'ingresso $x_I[n]$, le risposte impulsive di canale $h_i[k]$ e i segnali d'uscita $x_O^{(i)}[n]$.

I metodi alternativi per la stima CAP già esistenti, come ricordato nel paragrafo 1.2.2, presentano limitazioni di tipo diverso, ma sempre tali da renderne generalmente svantaggioso o impraticabile l'uso. Nell'insieme di tali metodi, quello proposto da Hikichi e Miyoshi [13, 14, 50], pur presentando criticità molto più accentuate dell'HMK rispetto alle dimensioni del problema (cfr. punti 1 e 2 dell'elenco qui sopra), sembra fornire un buon punto di partenza rispetto ai punti 3 e 4. Tale metodo, di grande innovazione formale, fonda la ricerca del polinomio dei poli comuni sul calcolo del polinomio caratteristico di una nuova *matrice di predizione lineare*.

Proprio la predizione lineare sembra rappresentare potenzialmente un ingrediente importante per la stima di una componente AR, comune a più segnali, perché di per sé è un metodo:

- Non supervisionato (cfr. punto 4);
- Basato sull'uso dei segnali (cfr. punto 3);
- Per cui il nutrito retroterra offerto dalla letteratura [23, 51, 52, 53, 54] offre metodi che non risentono, in sé, della lunghezza L_H delle $h_i[k]$ [55].

Il metodo di Hikichi e Miyoshi, pur formalmente molto elegante, per ammissione degli stessi autori [13] mostra una scarsa robustezza computazionale al crescere dell'ordine dei CAP (cfr. punto 1) e del numero dei canali.

4.2 Nuove idee intorno alla predizione lineare

L'idea di base intorno a una generica architettura per la stima dei poli comuni, basata sulla conoscenza dei soli canali di uscita del sistema acustico, può riassumersi schematicamente come in figura 4.2. Limitandosi sempre alla più

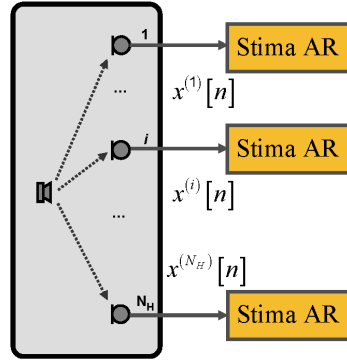


Figura 4.2: Uno schema generico di sistema per la stima di CAP: N_H stime AR in parallelo, operanti su diversi canali (e quindi segnali), ma alle quali è richiesto di produrre lo stesso risultato, ovvero il denominatore dei poli comuni.

semplice configurazione SIMO, tale schema riassume il nucleo centrale da cui prendono le mosse le idee di questo capitolo: disporre di un algoritmo in grado di operare stime AR su ognuno dei canali di uscita del sistema acustico, e allo stesso tempo in grado di imporre che tali stime (ognuna operata su un canale diverso, e quindi su un diverso segnale) generino lo stesso risultato, ossia forniscano proprio il denominatore comune di cui i CAP sono le radici.

La predizione lineare si presta naturalmente ad essere scelta come il principale ingrediente metodologico per la stima dei coefficienti del modello AR. Il generico schema che ne illustra i presupposti per un problema a singolo canale è quello di figura 4.3. Ammettendo

$$A(z) = 1 + \alpha(z) \quad \text{con} \quad \alpha(z) = a_1 z^{-1} + \dots + a_{L_A} z^{-L_A} \quad (4.2)$$

e

$$C(z) = c_1 z^{-1} + \dots + c_P z^{-P} \quad (4.3)$$

se $P = L_A$ la minimizzazione dell'energia del segnale $e[n]$ è raggiunta per $C(z) = -\alpha(z)$, ovvero $c_i = -a_i \forall i = 1, \dots, L_A$. $C(z)$, filtro non ricorsivo privo di propagazione istantanea, è detto filtro di Predizione Lineare (o LP dall'ovvia permutazione dei vocaboli anglofona), e P ordine di predizione lineare. I coefficienti del filtro LP sono determinati proprio in base alla minimizzazione dell'energia dell'errore di predizione lineare $e[n]$ in senso statistico, e in tali

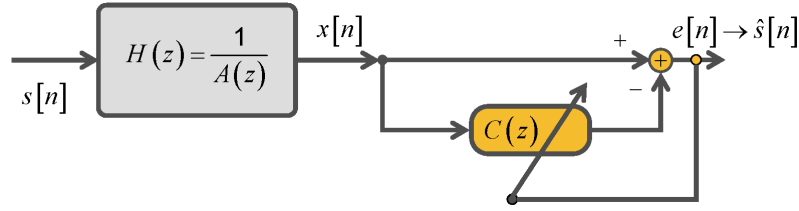


Figura 4.3: Un sistema puramente AR, $H(z)$, seguito da una generica architettura di predizione lineare: la ricerca dei coefficienti del sistema $C(z)$, non ricorsivo e privo del ritardo nullo, attraverso la minimizzazione dell'energia del segnale di errore $e[n]$, porta $C(z)$ a compensare $A(z)$ e $e[n]$ a fornire una stima del segnale $s[n]$.

condizioni si può assumere $e[n] = s[n]$. Il filtro LP punta dunque a compensare gli effetti di $1/A(z)$, e in tal modo fornisce una stima del denominatore $A(z)$, a meno di un segno e dell'aggiunta dell'unità.

Nel caso di un sistema acustico SIMO come quello in figura 4.1 la questione è in realtà più complessa, perché ognuno degli N_H canali $H_i(z) = B^{(i)}(z)/A(z)$ non è un puro sistema AR, ma è dotato di un numeratore $B^{(i)} \neq 1$, e rientra per questo nella più generica categoria dei sistemi ARMA. In tali casi, un semplice sistema a predizione lineare come quello in figura 4.3 non è in generale sufficiente a stimare $A(z)$, e delle modifiche si rendono necessarie rispetto ai normali metodi basati sulla semplice LP [56, 57]. A titolo di esempio, il celebre metodo di Yule Walker per il calcolo del filtro LP e dell'associata stima spettrale è da sostituire con il *Modified* Yule Walker quando il sistema $H(z)$ è ARMA [23]. Il realtà, in un contributo del 2003 nel contesto della deconvoluzione acustica per segnali vocali, Gaubitch e altri [58] mostrano che quando la predizione lineare è calcolata su più canali, anche un semplice procedimento di media in senso spaziale permette di recuperare una stima sostanzialmente accurata della componente AR comune, nonostante la presenza delle diverse componenti MA nei rispettivi e diversi canali. Nel contesto di un sistema multicanale come quello di figura 4.2, dunque, l'uso metodi basati sulla semplice predizione lineare può essere ancora considerato un valido approccio per la stima del denominatore comune, mantenendo in più il vantaggio della semplicità, utile nell'estensione dei metodi stessi alla multicanalità condizionata.

Nell'ottica di ideare un nuovo metodo per la stima del denominatore dei CAP basato sulla predizione lineare, è necessario innanzi tutto individuare un buon algoritmo di stima del filtro LP tale da:

1. Offrire prestazioni adeguate nella stima;
2. Essere facilmente estendibile al caso multicanale

L'idea dell'estensione punta alla costruzione di un algoritmo da battezzare ragionevolmente con il nome di *Predizione Lineare Multicanale Vincolata* sulla base del principio già introdotto: realizzare un'istanza dell'algoritmo LP per ognuno degli N_H canali, con il vincolo che il risultato prodotto sia il medesimo in tutti i canali.

Graficamente, l'idea appena esposta è rappresentata in figura 4.4.

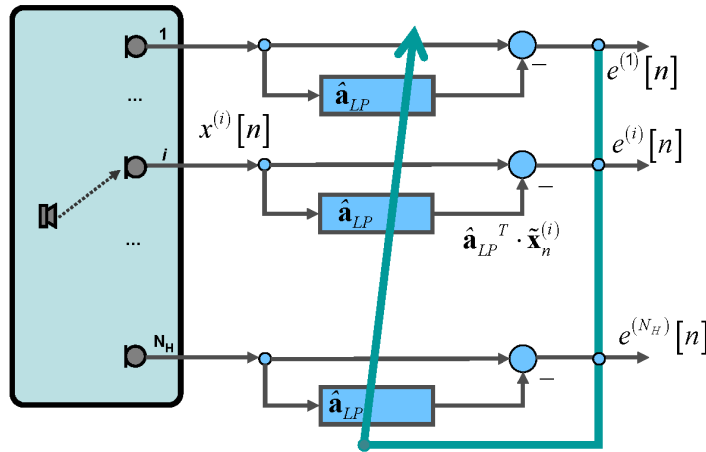


Figura 4.4: Una generica struttura per la Predizione Lineare Multicanale Vincolata: un algoritmo punta a minimizzare l'energia degli errori di predizione in tutti i rispettivi N_H canali, con il vincolo che il filtro di predizione lineare \mathbf{a}_{LP} sia sempre il medesimo.

Per ciò che riguarda la scelta dell'algoritmo di singolo canale da cui partire, una prima grande differenza tra i metodi rintracciabili in letteratura è la struttura di riferimento del filtro LP: due possibili e frequenti scelte sono costituite, da una parte, da una struttura trasversale, o di tipo FIR, e dall'altra da un filtro a traliccio o, in inglese, *Lattice*. Una seconda differenza riguarda il tipo di stima: da una parte un approccio adattativo, con i coefficienti del filtro aggiornati adattativamente e in diretta come funzione del segnale $x[n]$; dall'altra un approccio di blocco. In quest'ultimo caso i valori dei filtri sono calcolati a partire da grandezze stimate dall'insieme dei valori assunti dal segnale $x[n]$ in tutta la sua evoluzione dinamica, oppure nell'arco di un'ampia finestra temporale. Dall'incrocio delle diverse possibilità, rispettivamente per la struttura del filtro e per la tipologia di algoritmo, nascono le seguenti quattro direzioni percorribili:

- Algoritmo **adattativo** con filtro **trasversale**. La predizione lineare è un problema per cui questo approccio è in realtà generalmente evitato, per due principali motivi:
 - Il segnale $x[n]$ diventa l'ingresso del filtro adattativo e quindi quello da cui va generata la matrice di correlazione \mathbf{R} , il cui condizionamento regola le prestazioni dell'algoritmo. Se anche il segnale sorgente $s[n]$ ha uno spettro $S(f)$ sostanzialmente piatto, per definizione $X(f) = S(f)/A(z)$ esibirà una certa dinamica; ciò rende comunemente alto $\text{cond}(\mathbf{R})$, con conseguenti problemi nella convergenza del filtro.
 - Il segnale di errore $e[n]$ è pari a $s[n]$ nella configurazione ad energia minima. Anche in queste condizioni ideali la non nullità di $e[n]$ induce delle variazioni adattative non nulle nel filtro adattativo LP anche quando questo ha raggiunto il suo valore ottimo. Per questo motivo la convergenza è sempre da intendersi da un punto di vista statistico, e non fornisce mai, all'equilibrio, dei valori stabili per il filtro LP.

Percorrere questa strategia significa in ogni caso dover optare per algoritmi adattativi estremamente robusti rispetto a $\text{cond}(\mathbf{R})$, e nessuna alternativa al complesso RLS sembra godere di tali proprietà. D'altra parte, la complessità computazionale dell'RLS cresce secondo la seconda potenza della lunghezza dei filtri [23], e ciò ne fa un candidato indesiderabile alla luce delle premesse fatte nel paragrafo 4.1.

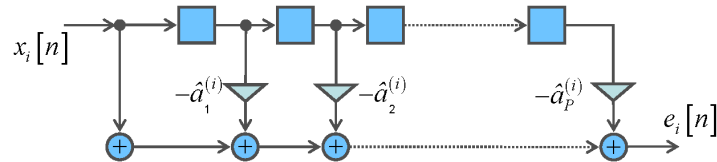


Figura 4.5: La tipica realizzazione diretta di un filtro trasversale, usato per la predizione lineare. Il formalismo usato per i coefficienti del filtro LP rispecchiano la corrispondenza con i coefficienti AR stimati: la cuspide indica una stima del corrispondente simbolo.

- Algoritmo **adattativo** con filtro **a traliccio**. Il filtro a traliccio, riportato in figura 4.6, deve la sua origine proprio al calcolo della predizione lineare, e di conseguenza ben si presta a questa applicazione. I coefficienti di riflessione γ_i (con $i = 1, \dots, P$) sono determinati in modo minimizzare congiuntamente l'energia dei due errori di predizione $f_P[n]$ e $b_P[n]$, detti

rispettivamente *in avanti* e *all'indietro*. Il più semplice algoritmo adattativo per questa struttura è il GAL [59, 23], dall'inglese Gradient Adaptive Lattice. Come suggerisce il nome, è basato sul metodo del gradiente similmente all'LMS, ma in questo caso la particolare struttura a traliccio garantisce una convergenza disaccoppiata per ognuno dei coefficienti di riflessione [60].

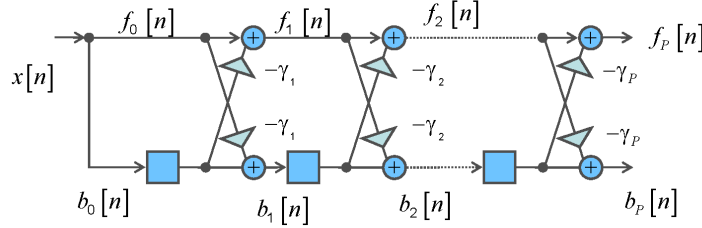


Figura 4.6: Un filtro a traliccio, o *lattice*. $f_k[n]$ e $b_k[n]$ sono, rispettivamente, gli errori di predizione in avanti (*forward*) e all'indietro (*backward*) di ordine k . I γ_k sono invece i coefficienti di riflessione di ordine k .

- Algoritmo **di blocco** con filtro **trasversale**. Gli algoritmi di blocco godono tradizionalmente di un uso più diffuso nella stima dei filtri LP. Il metodo di Yule-Walker sarebbe in questo caso l'algoritmo di riferimento.
- Algoritmo **di blocco** con filtro **a traliccio**. In questo caso l'algoritmo di riferimento è il metodo di Burg [61, 55, 19, 23].

Una sintesi schematica delle possibilità appena elencate è presentata in tabella 4.1.

Tabella 4.1: Sintesi dei quattro possibili algoritmi selezionabili per la Predizione Lineare Multicanale Vincolata, sulla base della struttura del filtro adattativo (diversa colonna) e della tipologia di algoritmo (diversa riga).

	Trasversale	Lattice
Adattativo	RLS	GAL
Di blocco	Yule-Walker	Burg

E' bene anticipare che delle prove informali sono state condotte per esplorare la percorribilità della prima, seconda e quarta ipotesi. Anche sorvolando sull'intrinseco peso computazionale, l'RLS, anche nella sua forma esponenzialmente pesata [23], presenta problemi di robustezza per P crescente, e necessita

di accurate strategie di regolarizzazione per scongiurare la divergenza al crescere dell'ordine. Sono necessarie inoltre ulteriori accortezze per minimizzare l'incidenza negativa delle componenti MA sulla minimizzazione dell'energia dei segnali di errore. Per l'eccessivo peso computazionale, la scarsa robustezza e le troppe complicazioni realizzative questo approccio è stato scartato.

Sempre nell'ambito dei metodi adattativi, il GAL ha prodotto risultati incoraggianti, ma al crescere dell'ordine di predizione i tempi di adattamento associati ai coefficienti “più lenti” crescono secondo dipendenze $> O(P)$. Per questa ragione anche questa strada è stata abbandonata.

Vere e proprie prove informali degne di nota non sono state condotte per il metodo di Yule-Walker; d'altro canto il metodo di Burg, fondato sulla robusta architettura a traliccio, ha permesso di raggiungere risultati promettenti mantenendo una relativa semplicità realizzativa. Il paragrafo seguente, oltre a ricordare i fondamenti dell'algoritmo di Burg per i problemi di stima AR a singolo canale, spiegherà come quest'ultimo è stato esteso nel senso della Predizione Lineare Multicanale Vincolata.

4.3 Predizione Lineare Multicanale Vincolata e metodo di Burg

Il metodo di Burg punta alla stima del filtro di predizione lineare ottimo, nella già citata forma a traliccio di figura 4.6. Il filtro, di ordine P , è definito dal vettore dei coefficienti di riflessione

$$\gamma^{(P)} = [\gamma_1, \gamma_2, \dots, \gamma_P]$$

Esiste naturalmente una corrispondenza biunivoca tra tale vettore e il vettore $\mathbf{c} = [c_1, \dots, c_P]$ dei coefficienti del filtro trasversale di predizione lineare $C(z)$ (4.3). Una conveniente trasformazione è quella che permette di risalire direttamente al vettore $\hat{\mathbf{a}}$ dei coefficienti di $\hat{A}(z)$, stima del sistema AR cercato $A(z)$, ovvero

$$\hat{\mathbf{a}}^{(P)} = [\hat{a}_0^{(P)}, \hat{a}_1^{(P)}, \hat{a}_2^{(P)}, \dots, \hat{a}_P^{(P)}] = [1, -c_1, -c_2, \dots, -c_P] \quad (4.4)$$

La procedura ricorsiva che consente di passare da γ a $\hat{\mathbf{a}}^{(P)}$ è la seguente

$$\begin{cases} \hat{a}_j^{(j-1)} = 0 \\ \hat{a}_0^j = 1 \\ \hat{a}_i^{(j)} = \hat{a}_i^{(j-1)} [n-1] - \gamma_j \hat{a}_{j-i}^{(j-1)} [n] \end{cases} \quad (4.5)$$

Un esame delle (4.5) suggerisce che per la conoscenza di $\hat{\mathbf{a}}^{(P)}$ è necessaria non solo la conoscenza di $\gamma^{(P)}$, ma anche quella di $\hat{\mathbf{a}}^{(P-1)}, \dots, \hat{\mathbf{a}}^{(1)}$, ovvero dei completi vettori dei filtri di predizione lineare trasversali per gli ordini $1, \dots, P-1$.

All'atto pratico, anche questi ultimi sono calcolabili a partire da $\gamma^{(P)}$ in base alla (4.5), e quindi la (4.5) non è che una formula iper-ricorsiva per il calcolo di $\hat{\mathbf{a}}^{(P)}$ a partire da $\gamma^{(P)}$.

Chiarito il contesto, e ricordata la corrispondenza tra la rappresentazione a traliccio e quella trasversale, è bene riassumere brevemente l'algoritmo di Burg per il singolo canale. Il punto di partenza, come già anticipato, è la minimizzazione della somma delle energie dei due errori di predizione, rispettivamente in avanti e all'indietro, espressa statisticamente come

$$J = E \left\{ [f_{L_A} [n]]^2 + [b_{L_A} [n]]^2 \right\} \quad (4.6)$$

essendo E l'operatore di valore atteso in senso statistico. La derivazione del metodo passa per la stima delle (4.6) in termini dei segnali osservati. Omettendo i passaggi intermedi, ciò consente di arrivare a un'equazione per ognuno dei P coefficienti di riflessione

$$\gamma_k \cdot \sum_{\ell=k}^{P-1} \left([f_{k-1} [\ell]]^2 + [b_{k-1} [\ell-1]]^2 \right) = 2 \sum_{\ell=k}^{P-1} f_{k-1} [\ell] b_{k-1} [\ell-1] \quad (4.7)$$

Risolvendo rispetto a γ_k questa porta direttamente a

$$\gamma_k = \frac{2 \sum_{\ell=k}^{P-1} f_{k-1} [\ell] b_{k-1} [\ell-1]}{\sum_{\ell=k}^{P-1} \left([f_{k-1} [\ell]]^2 + [b_{k-1} [\ell-1]]^2 \right)} \quad (4.8)$$

La (4.8) va ancora una volta usata ricorsivamente, dopo aver calcolato gli andamenti nel tempo dei due errori di predizione fino all'ordine $k-1$ con le seguenti espressioni

$$\begin{cases} f_k [n] = f_{k-1} [n] - \gamma_k b_{k-1} [n-1] \\ b_k [n] = b_{k-1} [n-1] - \gamma_k f_{k-1} [n] \end{cases} \quad (4.9)$$

e ricordando che

$$f_0 [n] = b_0 [n] = x [n] \quad (4.10)$$

come del resto appare chiaro dallo schema di figura 4.6.

Nel caso multicanale l'insieme degli N_H filtri di predizione a traliccio produce N_H coppie di errori di predizione (in avanti e all'indietro) per ogni valore dell'indice k . Per $P = L_A$ La nuova somma delle energie da minimizzare è in questo caso data dalla seguente espressione

$$J = \sum_{i=1}^{N_H} J^{(i)} = \sum_{i=1}^{N_H} E \left\{ [f_{L_A}^{(i)} [n]]^2 + [b_{L_A}^{(i)} [n]]^2 \right\} \quad (4.11)$$

Ora però dovrà richiedersi il rispetto del vincolo più volte citato, ovvero

$$\gamma_k^{(i)} = \gamma_k \quad \forall i = 1 \dots N_H \quad (4.12)$$

dove $\gamma_k^{(i)}$ è il coefficiente di riflessione di ordine k per il filtro a traliccio dell' i -esimo canale.

La minimizzazione della (4.11) produce N_H soluzioni disaccoppiate per gli N_H filtri a traliccio, date da N_H diversi set di equazioni come le (4.7), ovvero per il canale i -esimo ($i = 1, \dots, N_H$)

$$\gamma_k^{(i)} \cdot \sum_{\ell=k}^{P-1} \left(\left[f_{k-1}^{(i)}[\ell] \right]^2 + \left[b_{k-1}^{(i)}[\ell-1] \right]^2 \right) = 2 \sum_{\ell=k}^{P-1} f_{k-1}^{(i)}[\ell] b_{k-1}^{(i)}[\ell-1] \quad (4.13)$$

Se le N_H versioni della (4.13) vengono sommate membro a membro, il vincolo richiesto dalla (4.12) trova una semplice realizzazione analitica, e fornisce un'equazione per il k -esimo coefficiente γ_k del filtro a traliccio comune data da

$$\begin{aligned} \gamma_k \cdot \sum_{i=1}^{N_H} \sum_{\ell=k}^{P-1} \left(\left[f_{k-1}^{(i)}[\ell] \right]^2 + \left[b_{k-1}^{(i)}[\ell-1] \right]^2 \right) &= \dots \\ \dots &= \sum_{i=1}^{N_H} 2 \sum_{\ell=k}^{P-1} f_{k-1}^{(i)}[\ell] b_{k-1}^{(i)}[\ell-1] \end{aligned} \quad (4.14)$$

Risolta per γ_k , la (4.14) conduce infine alla seguente espressione

$$\gamma_k = \frac{\sum_{i=1}^{N_H} 2 \sum_{\ell=k}^{P-1} f_{k-1}^{(i)}[\ell] b_{k-1}^{(i)}[\ell-1]}{\sum_{i=1}^{N_H} \left\{ \sum_{\ell=k}^{P-1} \left(\left[f_{k-1}^{(i)}[\ell] \right]^2 + \left[b_{k-1}^{(i)}[\ell-1] \right]^2 \right) \right\}} \quad (4.15)$$

da accoppiare sempre con le relazioni ricorsive

$$\begin{cases} f_k^{(i)}[n] = f_{k-1}^{(i)}[n] - \gamma_k b_{k-1}^{(i)}[n-1] \\ b_k^{(i)}[n] = b_{k-1}^{(i)}[n-1] - \gamma_k f_{k-1}^{(i)}[n] \end{cases} \quad \forall i = 1, \dots, N_H \quad (4.16)$$

La (4.15) e la (4.16) forniscono le relazioni che definiscono la Predizione Lineare Multicanale Vincolata attraverso l'estensione del metodo di Burg. Al meglio della conoscenza dell'autore di questo lavoro, questa derivazione è di per sé originale, e tanto più in quanto inserita nel contesto della stima del denominatore dei CAP, nel senso di figura 4.4.

Prima di proseguire, sembra opportuno riassumere a quali obiettivi strutturali, tra quelli elencati all'inizio del capitolo, questo nuovo metodo introdotto permette di dare una risposta.

1. La complessità computazionale è inferiore a $O(P)$, perché tutte le equazioni ricorsive per coefficienti di riflessione e errori di predizioni aggiungono solo un nuovo passo per ogni incremento dell'ordine P ; è inoltre semplicemente lineare rispetto al numero di canali usati;
2. Da punto di vista computazionale, lavorando sui segnali, è insensibile rispetto alla complessità riverberante dell'ambiente scelto, nel senso che la lunghezza delle risposte impulsive $h_i[k]$ (cfr. fig. 4.1) non entra in gioco come fattore di criticità computazionale;
3. Non è necessaria la conoscenza delle risposte impulsive, quindi viene eliminato un livello di complessità architetturale rispetto al metodo HMK;
4. Non è richiesta la conoscenza del segnale usato per sollecitare i canali, e in questo senso il nuovo metodo, seppur costruito in modo assai semplice e con l'uso di tecniche tradizionali, è da considerarsi di tipo *Blind*, o non supervisionato.

4.4 Simulazioni e risultati sperimentali

4.4.1 Considerazioni preliminari

E' stato osservato alla fine dello scorso paragrafo come, per quanto riguarda le questioni realizzative e la dipendenza della complessità della dimensionalità del problema, il metodo proposto già risponde positivamente a molte delle questioni aperte elencate all'inizio di questo capitolo, nel paragrafo 4.1.

Resta da provare che l'accuratezza raggiunta da tale metodo nella stima del denominatore dei CAP fornisce livelli accettabili, e soprattutto valutare come, da tale punto di vista, il metodo proposto è comparabile con i metodi esistenti. Nello specifico, è sembrato opportuno usare come riferimento il metodo HMK già ampiamente citato, e ricordato in particolare nel paragrafo 1.2.1.

Le prove portate a termine sono fortemente debitorie dei risultati raggiunti alla fine del capitolo 3. Come si ricorderà, è stato lì formalizzato un modello semi-analitico basato sul paradigma dei CAP, in grado di:

1. Generare modelli verosimili di canali acustici basati su CAP, per semplici ambienti parallelepipedali con pareti dalle proprietà riflettenti indipendenti dalla frequenza, aventi una sorgente coincidente con uno dei vertici del parallelepipedo e sensori disposti a piacere;
2. Rendere disponibile la conoscenza analitica del denominatore dei poli comuni $A(z)$ associato ai canali dell'ambiente simulato.

Al meglio della conoscenza di chi scrive, in letteratura non esiste un metodo di generazione di modelli di canali acustici che goda delle proprietà appena evidenziate. Questo contesto offre quindi un'occasione unica non solo per provare l'efficacia del sistema di stima qui proposto, ma anche e soprattutto per mettere alla prova il metodo HMK in quanto tale; quest'ultimo è quindi usato non solo come metodo di riferimento, ma come vero e proprio oggetto della validazione. I risultati raggiunti hanno quindi in questo senso un doppio carattere di novità. Viene mostrato come non solo il nuovo metodo sia in grado di produrre stime più accurate del metodo HMK, ma anche come quest'ultimo, generalmente considerato come un metodo di riferimento, cessi di fornire risultati accurati già da ordini decisamente bassi.

4.4.2 Descrizione del set-up e delle prove

Il set-up virtuale usato per produrre modelli di canale risponde naturalmente alla descrizione generale data al punto 1 dell'elenco dello scorso paragrafo: il fine è produrre un sistema SIMO come quello genericamente descritto in figura 4.1, con un numero di canali $N_H = 32$. Il coefficiente di riflessione in ampiezza è stato scelto pari a 0.7 (corrispondente a un assorbimento in potenza pari a circa 0.51) le dimensioni pari a $(2 \times 2.5 \times 3) \text{ m}^3$ e le posizioni dei sensori scelte in modo casuale, con probabilità uniforme lungo ognuna delle tre dimensioni ortogonali della stanza.

La sequenza di operazioni eseguite può riassumersi come segue:

1. Nei termini descritti nel paragrafo 3.4 e per un assegnato valore dell'ordine L_A , è stato calcolato il denominatore comune del sistema, $A(z)$, e, per ogni accoppiamento tra la sorgente nel vertice e ognuno dei 32 sensori, è stato valutato un numeratore $B_i(z)$.
2. Delle $N_H = 32$ funzioni di trasferimento di canale $H_i(z) = B_i(z)/A(z)$ si è calcolata la risposta impulsiva $h_i[k]$.
3. Le 32 risposte impulsive sono state troncate al massimo dei T_{60} , ovvero poste a zero per $n > \bar{n}$, dove \bar{n} è definito da:

$$\bar{n} : \quad \text{primo valore per cui} \quad 10 \cdot \log_{10} \frac{h_i^2[\bar{n}]}{\max(h_i^2[n])} < -60 \text{ dB}$$

4. Le risposte impulsive troncate sono state usate con il metodo HMK, insieme ai veri ordini L_A ed L_B associati al modello generato. Sembra importante notare che un troncamento al T_{60} appare una condizione insolitamente ottimistica per le risposte impulsive su cui si basa il metodo HMK.

5. E' stata usata una sorgente di rumore bianco gaussiano, filtrata dai modelli di canale integri $H_i(z)$, per produrre $N_H = 32$ segnali filtrati $x_i[n]$ ($i = 1, \dots, 32$). Il metodo proposto è stato usato sui segnali $x_i[n]$, con un ordine di predizione P anche in questo caso esattamente uguale all'ordine del denominatore L_A del modello generato.

La procedura descritta, dal punto 1 al punto 4, è stata ripetuta al variare dell'ordine del denominatore comune, L_A , generato al punto 1: L_A è stato fatto variare da 6 a 30 per tutti i valori pari.

Per ogni valore di L_A , è stato prodotto:

- Un grafico contenente gli andamenti di $20 \cdot \log_{10}|1/A(z)|$, con $A(z)$ descritto dai tre seguenti rispettivi set di valori:
 - Il valore “vero” del denominatore comune di ordine L_A comune generato dal modello;
 - Il denominatore di ordine L_A stimato con il metodo HMK;
 - Il denominatore di ordine L_A stimato con il metodo proposto.
- Per ognuno dei due denominatori stimati, è stata calcolata una norma in grado di valutare quantitativamente la distanza dal profilo generato dal denominatore vero, calcolata come:

$$\delta_\alpha = \frac{2}{F_S} \int_0^{F_S/2} 20 \cdot \log_{10} \left| \frac{\hat{A}_\alpha(z)}{A(z)} \right| \cdot df \quad (4.17)$$

dove il pedice α assume separatamente i due valori *Burg* e *HMK*

4.4.3 Risultati

I risultati ottenuti sono raggruppati nelle tre figure 4.7, 4.8 e 4.9. La tabella 4.2 e la figura 4.10 contengono invece la sintesi dell'andamento di δ_{Burg} e δ_{HMK} al variare dell'ordine del denominatore L_A .

Dall'esame dei grafici di singola simulazione e dalle sintesi finali emergono due risultati molto importanti:

- Il metodo proposto permette di raggiungere un'accuratezza differenziale di pochi dB, sostanzialmente costante entro l'intervallo degli ordini considerati per L_A e in media inferiore a 1.5 dB. In ragione di tali risultati fornisce un'accuratezza molto maggiore del metodo HMK, già a partire da $L_A = 12$.
- Se si escludono i pochi primi ordini generati per L_A , le prove condotte mostrano che il metodo HMK ha delle prestazioni negativamente inattese

Tabella 4.2: Andamento della norma differenziale δ_α definita in (4.17), per $\alpha = HMK$ e $\alpha = Burg$, al variare dell'ordine L_A . Questi stessi risultati sono presentati graficamente in figura 4.10.

$P = L_A$	δ_{HMK} (dB)	δ_{Burg} (dB)	$P = L_A$	δ_{HMK} (dB)	δ_{Burg} (dB)
6	0	1.0	20	8.6	2.7
8	0.13	0.76	22	10.1	0.35
10	1.3	2.7	24	11.3	0.82
12	3.5	3.4	26	12.3	2.1
14	4.9	0.84	28	13.8	1.1
16	6.4	0.67	30	17.0	1.3
18	8.0	0.78			

nella stima del denominatore dei poli comuni, e anzi, al crescere dell'ordine del denominatore dei CAP, il parametro δ_{HMK} definito dalla (4.17) sembra crescere in modo quasi lineare con L_A . La scoperta dell'entità dei limiti del metodo HMK è, come anticipato, un'altra prerogativa di novità dei risultati riassunti in quest'ultimo capitolo, anche se in modo sostanzialmente inatteso. In passato l'efficacia del metodo era stata genericamente misurata in base alla capacità di descrivere picchi di risonanza in posizioni qualitativamente corrispondenti a quelle attese [6]. La definizione del modello semi-analitico nel capitolo 3 ha invece permesso qui di generare prove di validazione più realistiche, con $A(z)$ noto e di ordine comunque superiore a quelle poche unità, con cui l'HMK tende ancora a comportarsi bene.

4.5 Conclusioni

In quest'ultimo capitolo è stato introdotto un nuovo metodo per la stima del denominatore dei CAP battezzato come Predizione Lineare Multicanale Condizionata, e ottenuto a partire dal metodo di stima AR di Burg basato sul filtro di predizione a traliccio. Il nuovo modello semi-analitico di canali acustici basato su CAP introdotto nel capitolo 3 ha permesso di condurre delle prove quantitative dettagliate mirate al confronto tra il nuovo metodo e il metodo di riferimento HMK. Le prove effettuate non solo hanno mostrato che il metodo proposto ha ottime prestazioni di accuratezza nella stima del denominatore comune, superiori al metodo HMK già per $L_A = 12$, ma in modo ancora più sorprendente ha permesso di scoprire che l'accuratezza offerta dal metodo HMK, ben lungi

dall'essere rassicurante, cresce rapidamente all'aumentare dell'ordine L_A .

Dal punto di vista dei problemi aperti sulla stima dei CAP, il nuovo metodo è in grado di rispondere a gran parte delle esigenze rimaste in sospeso. In particolare, rispetto al metodo HMK offre una complessità computazionale molto ridotta, non risente delle caratteristiche riverberanti dell'ambiente scelto (perché la sua complessità non dipende dalla lunghezza delle risposte impulsive-campione) ha unicamente bisogno dei segnali di uscita del sistema multicanale, e rappresenta un metodo non supervisionato. Per la particolare forma del semplice modello semi-analitico usato nelle prove sperimentali, ci si è limitati all'applicazione a un modello di tipo SIMO. Dal punto di vista dell'applicabilità del metodo, i risultati raggiunti non suggeriscono comunque nessuna limitazione rispetto all'applicabilità a sistemi MIMO.

D'altra parte, il metodo proposto mantiene anche dei limiti che sembra opportuno sottolineare. Il metodo di Burg a singolo canale risente di alcuni difetti noti in letteratura; sebbene un paio tra questi siano importanti solo se si dispone di segnali di lunghezza molto limitata, sussiste un problema di scarsa robustezza rispetto a una non esatta scelta dell'ordine. Tale criticità è tanto più rilevante nel contesto della stima dei CAP, per i problemi legati all'impraticabilità dell'ordine vero ricordati nel paragrafo 1.4. L'analisi della robustezza del metodo proposto rispetto a una scelta dell'ordine diversa da quello vero rimane un aspetto da approfondire, e insieme un altro importante parametro per meglio definire il confronto delle prestazioni rispetto al metodo HMK. Un'ultima limitazione che sembra importante sottolineare è la maggiore difficoltà, qui, di stabilire metriche come quelle definite dalla (1.21), e che rendono più agevole il compito di selezionare il miglior ordine efficace per la stima del denominatore dei CAP. La ricerca di nuove metriche per coniugare il metodo di stima proposto con il metodo di selezione dell'ordine introdotto nel paragrafo 1.4.2 costituisce un altro problema lasciato aperto dal presente lavoro e che appare di grande interesse per futuri sviluppi.

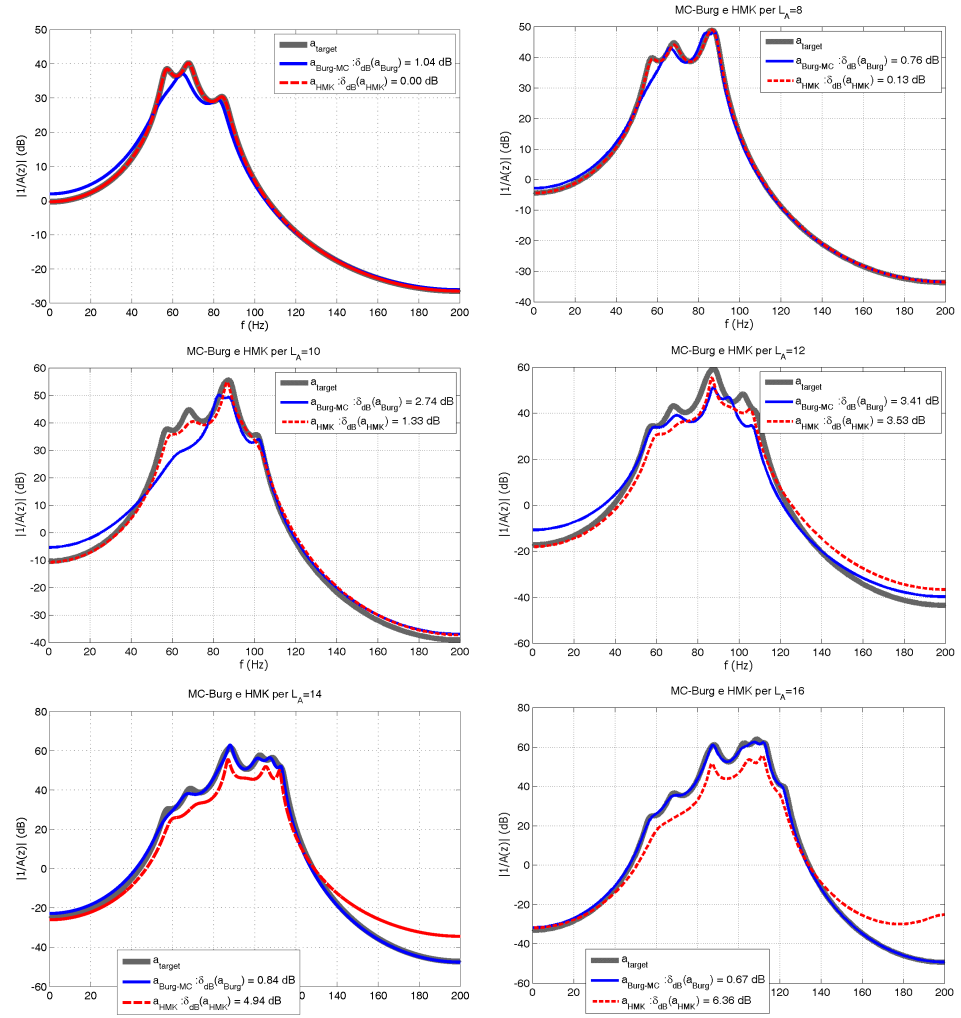


Figura 4.7: Rappresentazione spettrale del denominatore comune $A(z)$ prodotto dal modello (grigio), insieme alle due stime generate dai due ripetitivi metodi di stima usati: da una parte l'HMK (rosso), e dall'altra il metodo proposto (blu), per 6 diversi valori di L_A , da 6 a 16.

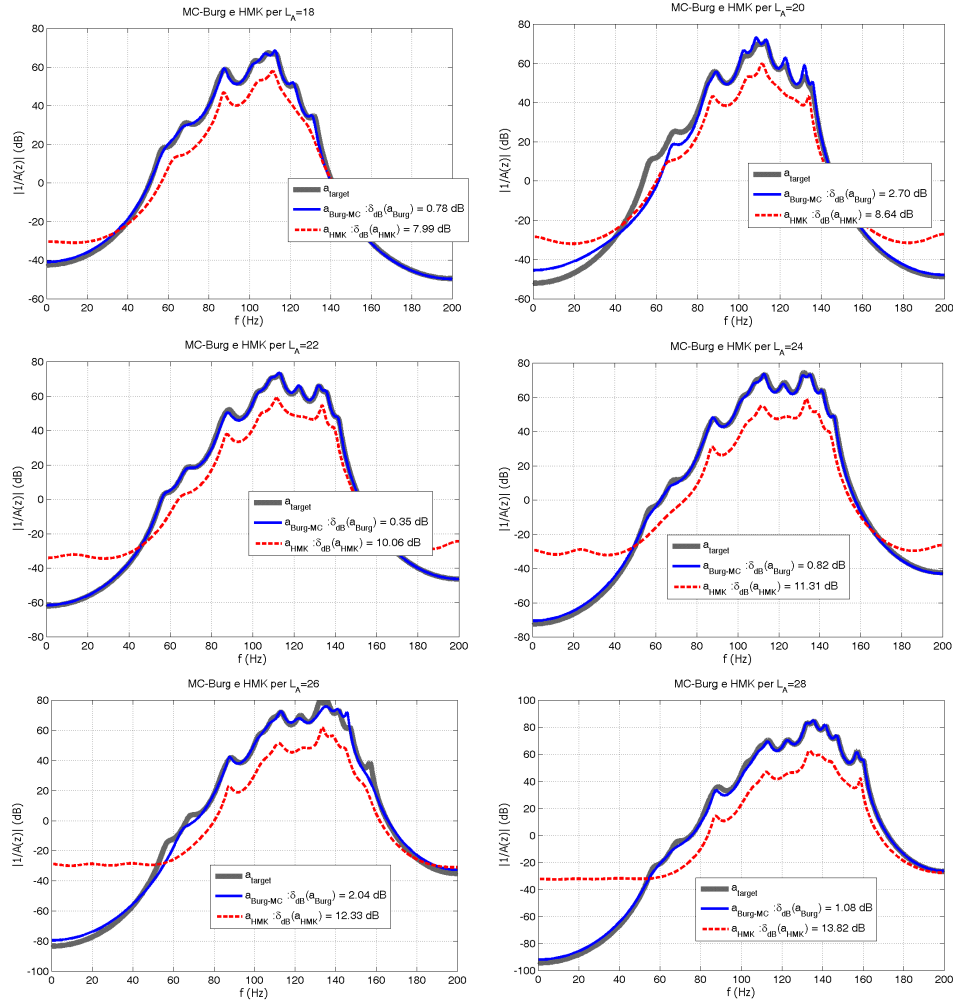


Figura 4.8: Rappresentazione spettrale del denominatore comune $A(z)$ prodotto dal modello (grigio), insieme alle due stime generate dai due ripetitivi metodi di stima usati: da una parte l'HMK (rosso), e dall'altra il metodo proposto (blu), per 6 diversi valori di L_A , da 18 a 28.

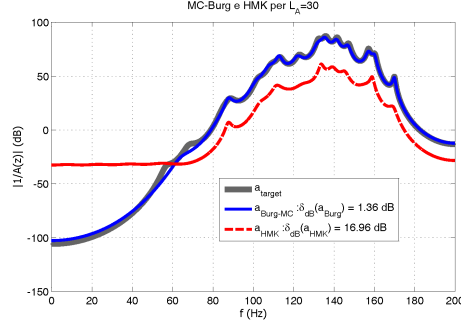


Figura 4.9: Rappresentazione spettrale del denominatore comune $A(z)$ prodotto dal modello (grigio), insieme alle due stime generate dai due ripetitivi metodi di stima usati: da una parte l'HMK (rosso), e dall'altra il metodo proposto (blu), per $L_A = 30$.

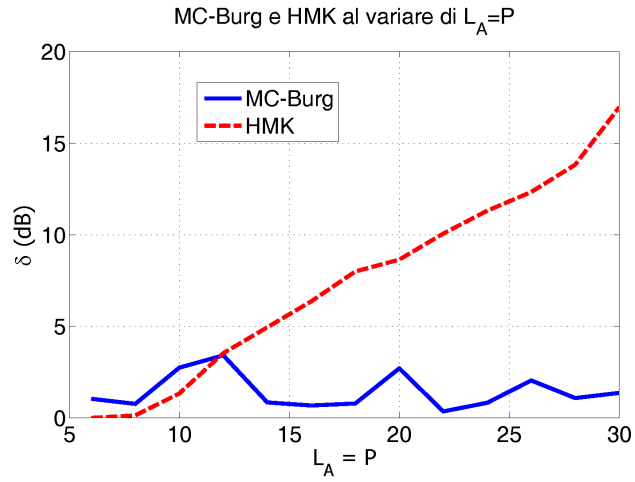


Figura 4.10: Andamento della norma differenziale δ_α definita in (4.17) per i due metodi usati nella stima del denominatore comune, al variare dell'ordine di quest'ultimo, L_A . Si osservi come, a partire da $L_A = 12$, il discostamento della stima prodotta dal metodo HMK continui a crescere, mentre quello del metodo proposto tenda a mantenersi stabilmente basso.

Bibliografia

- [1] Y. Haneda, S. Makino, and Y. Kaneda, “Modeling of a room transfer function using common acoustical poles,” in *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP-92)*, vol. 2, 23-26 March 1992, pp. 213 – 216. 1
- [2] R. Parisi, R. Gazzetta, and E. Di Claudio, “Prefiltering approaches for time delay estimation in reverberant environments,” in *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP '02)*, vol. 3, 13-17 May 2002, pp. 2997–3000. 1, 4
- [3] G. Bunkheila, A. Cirillo, M. Scarpiniti, R. Parisi, and A. Uncini, “Interfacce acustiche intelligenti in ambienti confinati,” in *Audio Engineering Society (AES) - Italian Section, Congresso annuale*, 2007. [Online]. Available: <http://www.aesitalia.org/Eventi/cng2007/programma.htm> 1
- [4] G. Bunkheila, R. Parisi, and A. Uncini, “Model order selection for estimation of common acoustical poles,” in *IEEE International Symposium on Circuits and Systems (ISCAS)*, 2008. 1, 15
- [5] G. Bunkheila, M. Scarpiniti, R. Parisi, and A. Uncini, “Stereo acoustical echo cancellation based on common poles,” in *IEEE International Symposium on Circuits and Systems (ISCAS)* (Submitted), 2009. 1
- [6] Y. Haneda, S. Makino, and Y. Kaneda, “Common acoustical pole and zero modeling of room transfer functions,” *IEEE Transactions on Speech and Audio Processing*, vol. 2, no. 2, pp. 320 – 328, April 1994. 3, 4, 5, 6, 13, 15, 29, 38, 39, 84
- [7] Y. Haneda, “Active noise control with a virtual microphone based on common-acoustical-pole and residue model,” in *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP '02)*, vol. 2, 2002, pp. 1877 – 1880. 4
- [8] Y. Haneda and Y. Kaneda, “Interpolation and extrapolation of room transfer functions based on common acoustical poles and their residues,” in

- IEEE ASSP Workshop on Applications of Signal Processing to Audio and Acoustics*, 19-22 Oct. 1997, p. Page(s):4 pp. 4
- [9] Y. Haneda, Y. Kaneda, and N. Kitawaki, "Common-acoustical-pole and residue model and its application to spatial interpolation and extrapolation of a room transfer function," *IEEE Transactions on Speech and Audio Processing*, vol. 7, no. 6, pp. 709–717, November 1999. 4
 - [10] Y. Haneda, S. Makino, and Y. Kaneda, "Common-acoustical-pole and zero modeling of head-related transfer functions," *IEEE Transactions on Speech and Audio Processing*, vol. 7, no. 2, pp. 188 – 196, March 1999. 4
 - [11] —, "Multiple-point equalization of room transfer functions by using common acoustical poles," *IEEE Transactions on Speech and Audio Processing*, vol. 5, no. 4, pp. 325 – 333, July 1997. 4
 - [12] F. Fontana, L. Gibin, D. Rocchesso, and O. Ballan, "Common pole equalization of small rooms using a two-step real-time digital equalizer," in *IEEE ASSP Workshop on Applications of Signal Processing to Audio and Acoustics*, 17-20 October 1999, pp. 195–198. 4
 - [13] T. Hikichi and M. Miyoshi, "Blind algorithm for calculating common poles based on linear prediction," in *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP '04)*, vol. 4, 17-21 May 2004, pp. iv–89 – iv–92. 5, 7, 72
 - [14] —, "A new algorithm for blind estimation of common poles in multiple transmission paths based on linear prediction," *Acoustical Science and Technology*, vol. 26, no. 1, pp. 1–7, 2005. 5, 7, 72
 - [15] P. Zuccarello, A. Gonzalez, J. Domingo, and G. Ayala, "Computation of common acoustical poles in subbands by means of a clustering technique," in *Proceedings of the 14th European Signal Processing Conference (EUSIPCO 2006)*, September, 4 - 8 2006. 5, 6
 - [16] G. Martinelli and M. Salerno, *Fondamenti di elettrotecnica*. La Siderea, 1995. 8, 9, 65
 - [17] L. Chua, C. Desoer, and E. Kuh, *Linear and nonlinear circuits*. McGraw-Hill New York, 1987. 8
 - [18] H. Kuttruff, *Room Acoustics (4th Ed.)*. Spoon Press, 2000. 9, 10, 11
 - [19] S. Kay and S. J. Marple, "Spectrum analysis - a modern perspective," *Proceedings of the IEEE*, vol. 69, no. 11, pp. 1380 – 1419, November 1981. 12, 70, 77

- [20] H. Akaike, "A new look at the statistical model identification," *Automatic Control, IEEE Transactions on*, vol. 19, no. 6, pp. 716–723, Dec 1974. 14
- [21] M. Karjalainen, P. Esquef, P. Antsalo, A. Mäkipvirta, and V. Välimäki, "Frequency-zooming arma modeling of resonant and reverberant systems," *Journal of the Audio Engineering Society*, vol. 50, no. 12, pp. 1012–1029, December 2002. 14
- [22] D. Campbell, K. Palomäki, and G. Brown, "A matlab simulation of shoebox room acoustics for use in research and teaching," *Computing and Information Systems*, vol. 9, no. 3, p. 48, 2005. [Online]. Available: <http://www.dcs.shef.ac.uk/~guy/pdf/campbell.pdf> 17
- [23] P. Clarkson, *Optimal and Adaptive Signal Processing*. CRC Press, 1993. 31, 33, 34, 42, 52, 72, 74, 76, 77
- [24] B. Farhang-Boroujeny, *Adaptive Filters: Theory and Applications*. John Wiley & Sons, 1998. 31, 42
- [25] S. Haykin, *Adaptive Filter Theory (2nd ed.)*. Prentice Hall, 1991. 31, 34, 42
- [26] D. G. Manolakis, V. K. Ingle, and S. M. Kogon, *Statistical and Adaptive Signal Processing: Spectral Estimation, Signal Modeling, Adaptive Filtering and Array Processing*. Artech House, April 2005. 31, 42
- [27] S. Sankaran and A. Beex, "Convergence behavior of affine projection algorithms," *Signal Processing, IEEE Transactions on [see also Acoustics, Speech, and Signal Processing, IEEE Transactions on]*, vol. 48, no. 4, pp. 1086–1096, April 2000. 31
- [28] S. Gay and S. Tavathia, "The fast affine projection algorithm," in *Acoustics, Speech, and Signal Processing, 1995. ICASSP-95., 1995 International Conference on*, vol. 5, 9-12 May 1995, pp. 3023–3026vol.5. 31, 52
- [29] A. V. Oppenheim and R. W. Schaffer, *Digital signal processing*. Prentice-Hall, 1975. 31
- [30] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge University Press, 1985. 33
- [31] J. Benesty, D. Morgan, and M. Sondhi, "A better understanding and an improved solution to the specific problems of stereophonic acoustic echo cancellation," *IEEE Transactions on Speech and Audio Processing*, vol. 6, no. 2, pp. 156–165, March 1998. 41, 42, 43, 44, 52

- [32] M. Sondhi, D. Morgan, and J. Hall, "Stereophonic acoustic echo cancellation-an overview of the fundamental problem," *Signal Processing Letters, IEEE*, vol. 2, no. 8, pp. 148–151, Aug. 1995. 42, 46
- [33] J. Benesty, D. Morgan, J. Hall, and M. Mohan Sondhi, "Stereophonic acoustic echo cancellation using nonlinear transformations and comb filtering," in *Acoustics, Speech, and Signal Processing, 1998. ICASSP '98. Proceedings of the 1998 IEEE International Conference on*, vol. 6, 12-15 May 1998, pp. 3673–3676. 42, 46, 47
- [34] A. Khong, J. Benesty, and P. Naylor, "Stereophonic acoustic echo cancellation: analysis of the misalignment in the frequency domain," *Signal Processing Letters, IEEE*, vol. 13, no. 1, pp. 33–36, Jan. 2006. 43, 44
- [35] S. Shimauchi and S. Makino, "Stereo projection echo canceller with true echo path estimation," in *Acoustics, Speech, and Signal Processing, 1995. ICASSP-95., 1995 International Conference on*, vol. 5, 9-12 May 1995, pp. 3059–3062vol.5. 46
- [36] J. Benesty, P. Duhamel, and Y. Grenier, "A multichannel affine projection algorithm with applications to multichannel acoustic echo cancellation," *Signal Processing Letters, IEEE*, vol. 3, no. 2, pp. 35–37, Feb. 1996. 46
- [37] J. Benesty, D. Morgan, and M. Sondhi, "A hybrid mono/stereo acoustic echo canceler," *Speech and Audio Processing, IEEE Transactions on*, vol. 6, no. 5, pp. 468–475, Sept. 1998. 47
- [38] V. Turbin, A. Gilloire, P. Scalart, and C. Beaugeant, "Adaptive filtering algorithms for stereophonic acoustic echo cancellation," in *Proceedings of the IWAENC*, September 1997, pp. 53–56. 47
- [39] A. Gilloire and V. Turbin, "Using auditory properties to improve the behaviour of stereophonic acoustic echo cancellers," in *Acoustics, Speech, and Signal Processing, 1998. ICASSP '98. Proceedings of the 1998 IEEE International Conference on*, vol. 6, 12-15 May 1998, pp. 3681–3684vol.6. 47
- [40] D. Morgan, J. Hall, and J. Benesty, "Investigation of several types of nonlinearities for use in stereo acoustic echo cancellation," *IEEE Transactions on Speech and Audio Processing*, vol. 9, no. 6, pp. 686–696, Sept. 2001. 47
- [41] S. L. Gray and J. Benesty, Eds., *Acoustic Signal Processing for Telecommunication*. Kluwer Academic Press, 2000. 47, 52
- [42] T. Gansler and J. Benesty, "New insights into the stereophonic acoustic echo cancellation problem and an adaptive nonlinearity solution," *IEEE*

- Transactions on Speech and Audio Processing*, vol. 10, no. 5, pp. 257–267, July 2002. 47, 48
- [43] J. Benesty, F. Amand, A. Gilloire, and Y. Grenier, “Adaptive filtering algorithms for stereophonic acoustic echo cancellation,” in *Acoustics, Speech, and Signal Processing, 1995. ICASSP-95., 1995 International Conference on*, vol. 5, 9–12 May 1995, pp. 3099–3102vol.5. 48
 - [44] P. Eneroth, S. Gay, T. Gansler, and J. Benesty, “A real-time implementation of a stereophonic acoustic echo canceler,” *Speech and Audio Processing, IEEE Transactions on*, vol. 9, no. 5, pp. 513–523, July 2001. 48
 - [45] H. Buchner, J. Benesty, and W. Kellermann, “Generalized multichannel frequency-domain adaptive filtering: Efficient realization and application to hands-free speech communication,” *Signal Processing*, vol. 85, no. 3, pp. 549–570, March 2005. 48
 - [46] Q. G. Liu, B. Champagne, and K. C. Ho, “On the use of a modified fast affine projection algorithm in subbands for acoustic echo cancellation,” in *Proceedings of the 7th IEEE Digital Signal Processing Workshop*, Norway, 1996, pp. 354–357. 48
 - [47] S. Chiucchi and F. Piazza, “A virtual stereo approach to stereophonic acoustic echo cancellation,” in *Circuits and Systems, 2000. Proceedings. ISCAS 2000 Geneva. The 2000 IEEE International Symposium on*, vol. 4, 28–31 May 2000, pp. 745–748vol.4. 48
 - [48] S. L. Gay, *Acoustic Signal Processing for Telecommunication*. Kluwer Academic Publishers, 2000, ch. The Fast Affine Projection Algorithm, pp. 23–41. 52
 - [49] E. Chau, H. Sheikhzadeh, and R. Brennan, “Complexity reduction and regularization of a fast affine projection algorithm for oversampled subband adaptive filters,” in *Acoustics, Speech, and Signal Processing, 2004. Proceedings. (ICASSP '04). IEEE International Conference on*, vol. 5, 17–21 May 2004, pp. V–109–12vol.5. 52
 - [50] M. Delcroix, T. Hikichi, and M. Miyoshi, “Blind dereverberation algorithm for speech signals based on multi-channel linear prediction,” *Acoustical Science and Technology*, vol. 26, no. 5, pp. 432–439, 2005. 72
 - [51] J. Makhoul, “Linear prediction: A tutorial review,” *Proceedings of the IEEE*, vol. 63, no. 4, pp. 561–580, 1975. 72
 - [52] —, “Spectral analysis of speech by linear prediction,” *Audio and Electroacoustics, IEEE Transactions on*, vol. 21, no. 3, pp. 140–148, Jun 1973. 72

- [53] L. Jackson, R. Rao, and S. Wood, "Parameter estimation by linear prediction in cascade form," in *Acoustics, Speech, and Signal Processing, IEEE International Conference on ICASSP '77*, vol. 2, May 1977, pp. 727 – 731. 72
- [54] S. Jackson, L.; Wood, "Linear prediction in cascade form," *Acoustics, Speech, and Signal Processing [see also IEEE Transactions on Signal Processing]*, *IEEE Transactions on*, vol. 26, no. 6, pp. 518 – 528, 1978. 72
- [55] D. Childers, Ed., *Modern Spectrum Analysis*. IEEE Press. 72, 77
- [56] S. Li and B. W. Dickinson, "An efficient method to compute consistent estimates of the ar parameters of an arma model," *IEEE Transactions on Automatic Control*, vol. AC-31, no. 3, pp. 275–278, March 1986. 74
- [57] —, "Application of the lattice filter to robust estimation of ar and arma models," *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. 4, no. 4, pp. 502–512, April 1988. 74
- [58] N. Gaubitch, P. Naylor, and D. Ward, "On the use of linear prediction for dereverberation of speech," in *Proceedings of the International Workshop on Acoustic Echo and Noise Control (IWAENC2003)*, Kyoto, September 2003, pp. 99–102. 74
- [59] L. Griffiths, "A continuously adaptive filter implemented as a lattice structure," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing*, 1977, pp. 683–686. 77
- [60] M. L. Honig and D. G. Messerschmitt, "Convergence properties of an adaptive digital lattice filter," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. ASSP-29, no. 3, pp. 642–653, June 1981. 77
- [61] J. Burg, "Maximum entropy spectral analysis," in *Proc. 37th Meeting of the Society of Exploration Geophysicist*, 1967. 77